Feedback Regulation in Gene Networks

by

Murat Acar

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author .....
Department of Physics
May 15, 2007

Certified by .....
Alexander van Oudenaarden
Associate Professor
Thesis Supervisor

Accepted by .....
Thomas J. Greytak
Chairman, Department Committee on Graduate Students

Feedback Regulation in Gene Networks

by

Murat Acar

Submitted to the Department of Physics
on May 15, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Cellular genetic information is encoded in DNA. The passage of this information from DNA to proteins is regulated at multiple levels and each level gives cells the chance to control the structure and function of their components. Transcriptional regulation is an important part of this multi-level process. By using feedback loops as part of transcriptional gene networks, cells can tune the level and stability of gene expression. In the first half of my thesis, I will discuss how isogenic cells can be re-programmed to have varying levels of memory associated with previous growth conditions and how genetic noise limits the stability of this cellular memory.

Noise in gene expression, through the phenotypic heterogeneity it promotes, has the potential to be a mechanism implemented by cells to cope with the uncertainties in environmental conditions. By randomly expressing multiple phenotypes, each fit to a certain environment, cells can survive unexpected changes in the extracellular environment. In this way, a population can hedge its bets against environmental uncertainty. Depending on how often the cells choose to display various phenotypes, the population can range from being highly diverse (heterogenenous) to being less diverse (homogenenous). In the second half of my thesis, I will discuss how the degree of phenotypic diversity for an isogenic population of cells can be tuned by re-engineering a gene network. I will present results from experiments which test the effect of noise-induced diversity on population fitness in the presence of fluctuating environments. The results demonstrate that for an optimum population growth in fluctuating environments, cells need to match the rates of inter-phenotypic switching to the frequency of environmental changes.

Thesis Supervisor: Alexander van Oudenaarden
Title: Associate Professor

Acknowledgements

I would like to thank my advisor Alexander van Oudenaarden for his mentorship and support, Ertugrul M. Ozbudak for introducing me to techniques in molecular biology and cloning, Attila Becskei for teaching me molecular biology techniques to genetically modify yeast cells, Mukund Thattai for insightful discussions on a variety of scientific and non-scientific topics, Jerome Mettetal for working with me on the project described in the third chapter of this thesis, Benjamin Kaufmann, Bernardo Pando, and Juan M. Pedraza for many useful discussions on various topics in quantitative biology, Mehran Kardar and Leonid Mirny for agreeing to be members of my thesis committee and my references, members of the van Oudenaarden group, Onur Ozcan, Serkan Cabi, and Can Kilic for all the fun during the past five years, and my family, my aunt, my uncle-in-law, and the rest of my extended family for their unconditional support and for raising me as a person who believes in the importance of education. Throughout my studies at MIT, I have been funded by an MIT Presidential Fellowship (endowed by Praecis Pharmaceuticals) awarded to myself and NIH and NSF grants awarded to my advisor.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Building complexity into cellular systems: one feedback loop at a time

Cells use gene and protein regulatory pathways to respond to changes in their intra- and extra-cellular environments. These changes can be in anything; from the concentrations of nutrients to temperature. Regulatory pathways are composed of functionally interacting pathway components, each specialized for a certain task in the cell. For example, for pathways responsible for metabolizing sugars, transporting the sugar into cells would be one such task. One big challenge awaiting the post-genomic era scientists is to find out how components of cellular pathways dynamically interact with each other to give rise to cellular structure and function. In this context, feedback regulation in gene networks arises as a mechanism implemented by cells to control gene expression levels. In the second chapter of my thesis, I will discuss how positive and negative feedback loops of a gene network take part in setting up the stability of gene expression levels in the network [5]. In the third chapter of my thesis, I will present results showing how feedback-loop mediated stability can be utilized as a mechanism for cells to cope with the uncertainties in the environmental conditions [6].

1.2 Central dogma of molecular biology and expression of gene network components

Every cell has a genetic identity established by the information encoded into its DNA. The transfer of this information from DNA to mRNA (via transcription) and then to proteins (via translation) is called the Central Dogma of Molecular Biology (Figure 1-1).

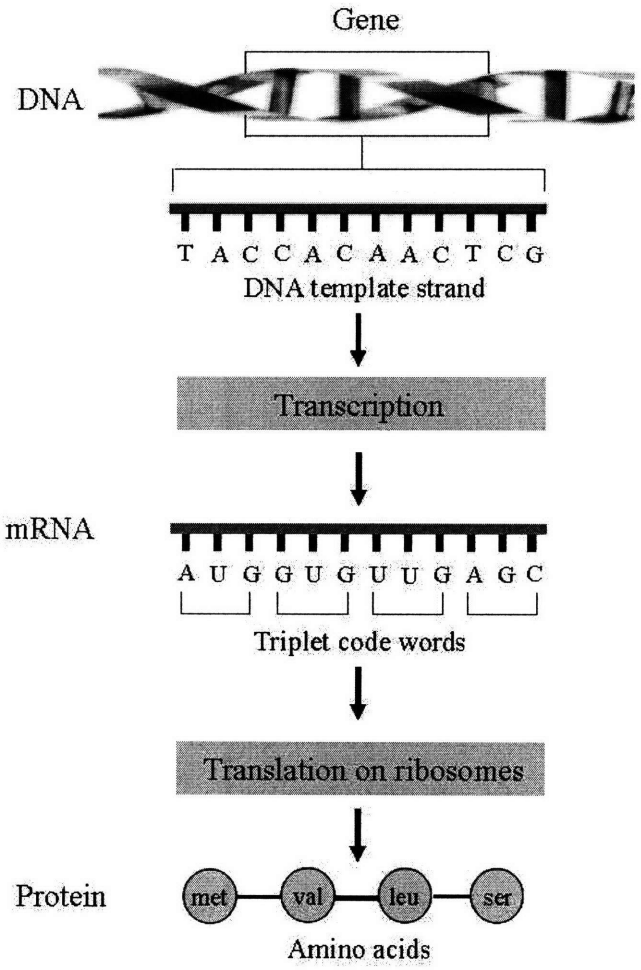

Figure 1-1: Flow of genetic information encoded in DNA to messenger RNA to protein. Figure adapted from [1]. Copyright ©Concepts of Genetics by William S. Klug and Michael R. Cummings. Reprinted by permission of Pearson Education, Inc.

DNA is a double stranded polymer composed of nucleotides as its monomers. Each nucleotide is composed of a base (adenine, thymine, guanine, or cytosine), a

sugar (deoxyribose), and phosphate groups. The two strands of the DNA are kept together through hydrogen bonds between bases on each strand. Adenine pairs up with thymine and guanine pairs up with cytosine. The number of hydrogen bonds formed between adenine-thymine and guanine-cytosine are two and three, respectively.

RNA, on the other hand, is a single stranded polymer. Like DNA, it is also composed of nucleotides as its monomers. However, there are a few differences between nucleotides of DNA and RNA. The nucleotides composing an RNA molecule carry uracil instead of thymine as one of the four bases and they have the sugar ribose instead of deoxyribose. There can be local double stranded structures on RNA (termed as 'RNA secondary structures') visualized as hairpin loops, because of the local base pair complementarities between adenine-uracil and guanine-cytosine.

Unlike DNA, there are various types of RNA, such as messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), micro RNA (miRNA), and small interfering RNA (siRNA). mRNA and tRNA are the major RNA types involved in the information passage from DNA to proteins.

An mRNA molecule with a sequence complementary to the DNA is synthesized (transcribed) from one of the strands of DNA by an enzyme called RNA polymerase. Upon transcription (and after the post-transcriptional modifications for eukaryotes), each mRNA goes to the ribosomes in the cytoplasm. Therefore, mRNAs deliver the genetic information from DNA to the ribosomes where translation occurs and the genetic information is encoded into the polypeptide chains. tRNAs play a role during translation; they attach aminoacids specified by the information carried by mRNA to the growing polypeptide chains.

The process of transcription involves three steps: initiation, elongation, and termination. It is initiated when RNA polymerase binds to the promoter region on DNA. Promoters are located upstream of genes on DNA. The DNA strand on which RNA polymerase lands is called the template strand. A primer is not necessary to initiate the transcription. Specific sequences on the promoter region are recognized by specific 'transcription factors' giving rise to the binding of those factors to the promoter. When this happens, the DNA becomes ready to attract RNA polymerase

enzyme to initiate the transcription from the promoter (Figure 1-2). Elongation in the 5'-to-3' direction follows the initiation and the process is terminated when the RNA polymerase runs into a termination sequence on the template DNA.

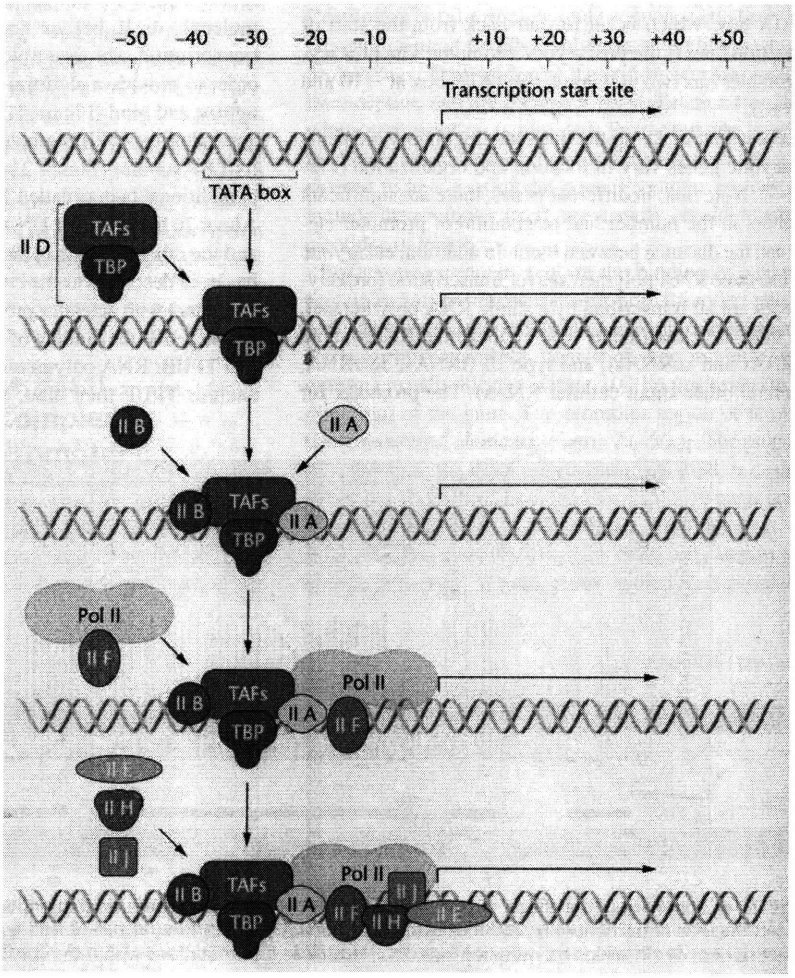

Figure 1-2: The assembly of transcription factors required for the initiation of transcription by RNA polymerase II [1]. Copyright ©Concepts of Genetics by William S. Klug and Michael R. Cummings. Reprinted by permission of Pearson Education, Inc.

The genetic code transcribed onto the mRNA molecules is translated to aminoacid chains at the ribosomal sites in the cytoplasm. Each 'word' within the genetic code (mRNA) is composed of three ribonucleotide 'letters'. These words are called 'codons' and each codon corresponds to a single aminoacid during the process of translation. Due to the fact that as many as 64 different triplet codons can be specified using a

four-letter alphabet and there are twenty amino acids present in cells, many amino acids are specified by more than one codon (Figure 1-3). In fact, this is the case for 18 out of 20 amino acids. The translation is initiated and terminated by certain codons. Most polypeptide chains start with the amino acid Methionine. The codon AUG encodes for this particular amino acid. The codons UAG, UAA, and UGA, on the other hand, are the stop codons terminating translation. They do not code for any amino acids. The aminoacid chains produced after translation are ultimately folded into proteins [1].

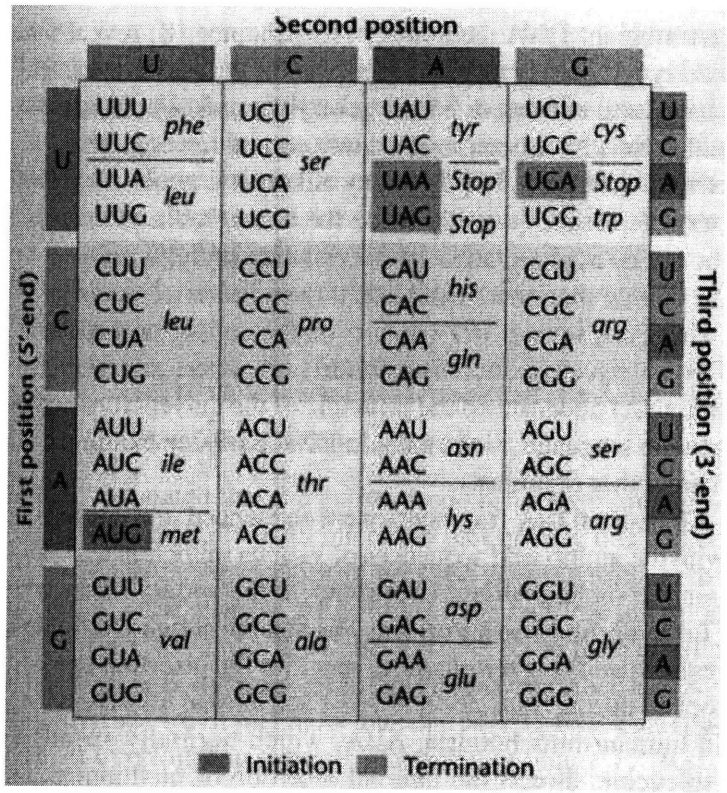

Figure 1-3: The coding dictionary. AUG encodes methionine, which initiates most polypeptide chains. All other amino acids except tryptophan, which is encoded only by UGG, are represented by two to six triplets. The triplets UAA, UAG, and UGA are termination signals and do not encode any amino acids [1]. Copyright ©Concepts of Genetics by William S. Klug and Michael R. Cummings. Reprinted by permission of Pearson Education, Inc.

In transferring the genetic information from DNA to the proteins, cells take an indirect approach by having an intermediate messenger molecule. In doing so, there

are at least two major advantages for cells: first, the more intermediate steps involved, the more opportunities obtained by cells in controlling the gene expression. Second, it is safer for the cell to have a molecule other than the DNA as the subject of biochemical 'processing' at the ribosomes.

1.3 Regulation of gene expression

Gene expression is a multi-step process including chromatin remodeling (specific to eukaryotes), transcription, post-transcriptional modifications (specific to eukaryotes), transport of mRNA from nucleus to cytoplasm (specific to eukaryotes), mRNA degradation, translation, and post-translational modifications (Figure 1-4).

Cells regulate the expression of their genes by controlling the efficiency and yield of these steps. By doing this, the amount of protein production in the cell is brought to specific levels.

Now if we look, in more detail, at some of the steps in gene expression regulation [1]:

Chromatin remodeling: Unlike prokaryotes, eukaryotes keep their DNA in a highly condensed form called 'chromatin'. DNA is tightly wrapped around special proteins such as histones. Unless the chromatin structure gets remodeled by opening up, the transcription factors and ultimately RNA polymerase enzyme cannot access the DNA to initiate the transcription (Figure 1-5). Therefore chromatin remodeling can be a very strong regulatory mechanism controlling gene expression. It occurs through the modification of histone tails by acetylation (adding an acetyl group), phosphorilation (adding a phosphate group), and methylation (adding a methyl group).

Post-transcriptional modifications: In eukaryotes, mRNAs (or 'pre-mRNAs' to be more precise) transcribed from the DNA are processed before their transport to the cytoplasm for translation. The processing of the 'pre-mRNA' includes capping at the 5' end to protect the mRNA from being degraded by the 5' exonuclease enzyme, splicing to get rid of noncoding introns from the mRNA, and addition of poly(A) tail to increase the half-life of the mRNA by giving the 3' exonuclease enzyme junk RNA

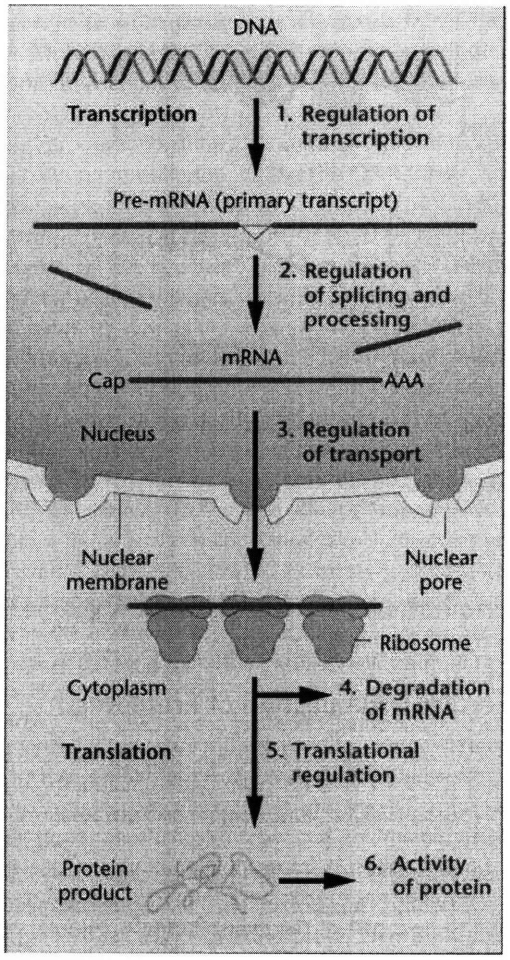

Figure 1-4: Various levels of regulation that are possible during the expression of the genetic material [1]. Copyright ©Concepts of Genetics by William S. Klug and Michael R. Cummings. Reprinted by permission of Pearson Education, Inc.

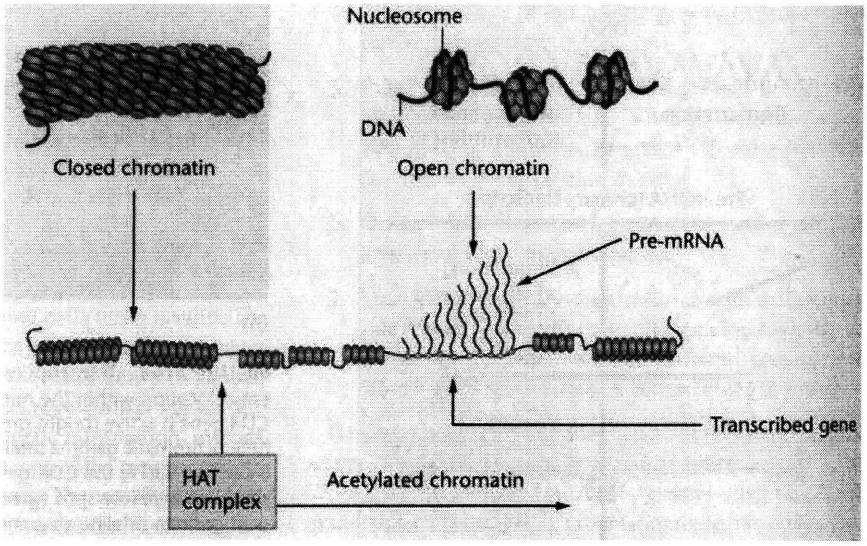

Figure 1-5: Histone acetyltransferases (HATs) are enzymes that add acetate groups to the tails of histones in nucleosomes of closed chromatin (top left). Action of a HAT complex converts closed chromatin to open chromatin (top right), allowing transription of mRNA to take place [1]. Copyright ©Concepts of Genetics by William S. Klug and Michael R. Cummings. Reprinted by permission of Pearson Education, Inc.

to 'chew'.

Post-translational modifications: After the translation, the structure and function of a protein can be modified through cleavage from one end or attachment of certain functional groups such as phosphate groups. All these manipulations are categorized as post-translational modifications. They give cells opportunities to control the amount, structure, and function of the proteins produced.

Although each cell in an organism carry the same genetic material and the same gene regulatory mechanisms are involved to produce proteins in cells, isogenic cells can gain non-genetic individuality by having differences in the number of proteins produced. Noise in gene expression is a mechanism giving rise to such individuality [2, 7, 8]. Noise can also enhance phenotypic diversity in isogenic populations by promoting cellular switching between the phenotypes expressed [5, 9].

1.4 Noise in gene expression

Noise in the expression of a single gene is defined as the standard deviation divided by the mean gene expression level [10].

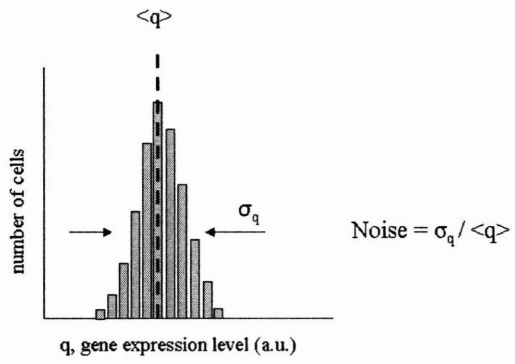

Figure 1-6: Variability or noise in the expression of a single gene for isogenic cells

Depending on its source, genetic noise can have intrinsic and extrinsic components [2].

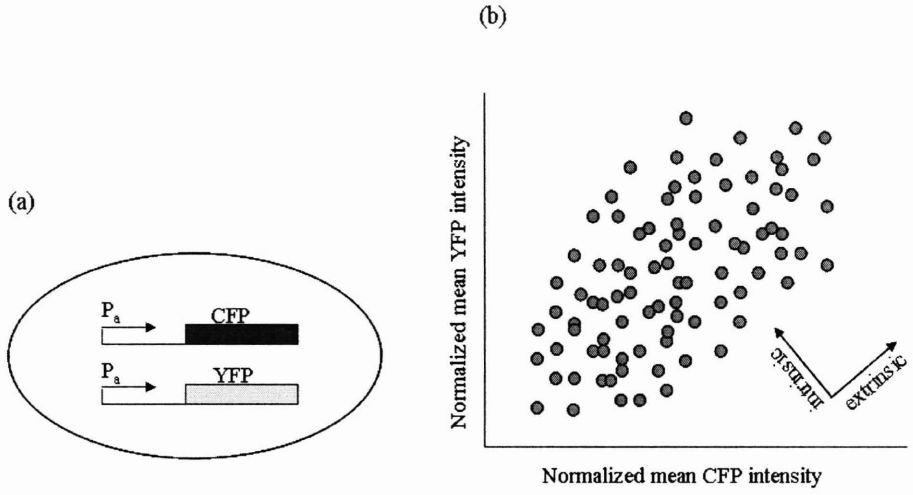

Figure 1-7: a, Cyan (CFP) and yellow (YFP) fluorescent proteins driven by the same promoter in single cells. b, Each filled-circle represents a single cell with mean CFP and YFP intensities. Variations perpendicular to the diagonal on which CFP and YFP intensities are equal correspond to intrinsic noise. However, cell-cell variations in parallel to the diagonal correspond to extrinsic noise. (Figure adapted from [2])

Intrinsic noise arises due to the inherently random nature of the biochemical processes such as mRNA decay. Intrinsic noise gives rise to differences even in the

expression levels of identical genes in a single cell. Extrinsic noise, on the other hand, arises due to effects which are the same for a single cell but change from one cell to another. Noise coming from differences in ribosome numbers, for example, is considered as an extrinsic noise source and generally termed as global noise affecting genes in a cell at the same level.

The physiological impact of the noise-induced non-genetic variability on cells is dependent on the number of molecules involved in the gene expression processes affected by noise. For example, the number of mRNA molecules in a cell is usually much lower compared to the number of proteins. Therefore, the physiological consequences of the stochastic variability in the number of mRNA molecules are expected to be much more drastic for cells than the consequences of the variability in the number of proteins.

Experiments [11, 12] designed to elucidate relative strengths of different noise sources in eukaryotes have shown that extrinsic noise is a more dominant noise source compared to the intrinsic noise. In addition to the progress made in understanding the origins of noise in gene expression, recent experimental studies [6, 13] have also addressed its biological relevance. For example, the work presented in the third chapter of this thesis has provided a direct experimental demonstration of how single cells can enhance survival in altering environmental conditions by utilizing stochasticity in their gene expression. Noise has the potential to be detrimental to the coordination of cellular activities; however, through the population heterogeneity it creates, noise might also act as a mechanism to help cells cope with the uncertainty in environments with fluctuating nutrient and fitness conditions [14–16]. In the third chapter of my thesis, I discuss how noise in gene expression can be beneficial for cells in constantly fluctuating environmental conditions. The results from this work give a quantitative basis for understanding the effect of gene expression noise on the growth of cell populations. As another example [13] to show how noise in gene expression can be useful for cells, Wernet and colleagues have shown that the stochastic expression of the transcription factor, spineless, defines the retinal mosaic structure necessary for color vision in the fruit fly *Drosophila*.

1.5 Structure and function of gene network motifs

When we look at the structure of gene regulatory pathways in cells, we realize that they are made up of recurrent network motifs. Network motifs are defined as patterns of connections between different network elements that occur much more frequently than they do in randomized networks [17]. For example, the feed-forward loop structure is a network motif. In the past few years, there has been an important amount of work [3, 18] done on identifying network motifs that are recurrent in regulatory networks of *E. coli* and *Saccharomyces cerevisiae*. For example, a genome-wide location analysis [3] has identified six different types of network motifs (Figure 1-8) employed by the yeast *Saccharomyces cerevisiae* to set its complex gene expression program. These include the autoregulatory motif, multi-component loop, the feedforward loop, the multi-input module, the single-input module, and the regulator cascade. These network motifs recur in a variety of different endogenous pathways [3] regulating gene expression in yeast.

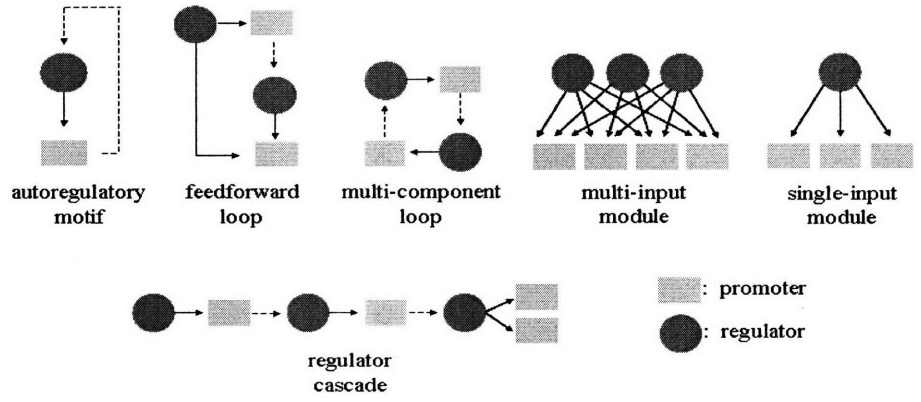

Figure 1-8: Network motifs in the yeast *Saccharomyces cerevisiae*. Light green rectangles and light blue circles represent promoters and regulators respectively. (Figure adapted from [3])

Network motifs can be considered as the building blocks of complexity in biological pathways. Therefore understanding the structure and function of these motifs is crucial for understanding complexity in cellular systems.

1.5.1 Autoregulatory motif

The autoregulatory motif is formed when a transcriptional regulator regulates its own expression by binding to the promoter of its own gene. The nature of this regulation can be either positive or negative. In the case of positive autoregulation, the transcriptional regulator enhances its own production. Perhaps the most important realization of positive autoregulatory systems is their ability to establish bistability [19]. However, positive autoregulation alone is not sufficient to give rise to bistability [20]. The system should also have enough nonlinearity. This could be achieved, for example, by the dimerization of the transcriptional regulator before binding to its own promoter. Negative autoregulation, or negative feedback, occurs when the transcriptional regulator represses its own expression level. Negative feedback loops provide systems with the ability to increase gene expression stability by decreasing expression level variations [21].

1.5.2 Multi-component loop

The multi-component loop is composed of two regulatory proteins regulating the expression level of one another. The nature of the regulation can be both positive and negative [20]. In the case of two negative regulatory loops, corresponding to a toggle-switch, the system switches between the ON or OFF states of both regulators in an alternating fashion. In the case of two positive regulatory loops corresponding to a positive feedback structure, on the other hand, the activity of the two regulators switches between their ON and OFF states in parallel.

1.5.3 Feedforward loop

In a feedforward loop motif (Figure 1-9) [4], two regulatory proteins (X and Y) independently bind to the promoter region of a gene (Z). One of the regulators (X) also controls the expression of the second regulator (Y).

Depending on the nature (activation or repression) of the interaction between the elements of this three-component structure, there can be 8 different (Figure 1-10)

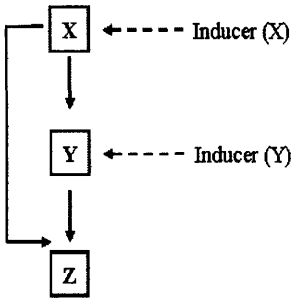

Figure 1-9: The structure of feedforward loop network motif. Inducer(X) and inducer(Y) activate or inhibit the activity of X and Y. (Figure adapted from [4])

feedforward motifs [4].

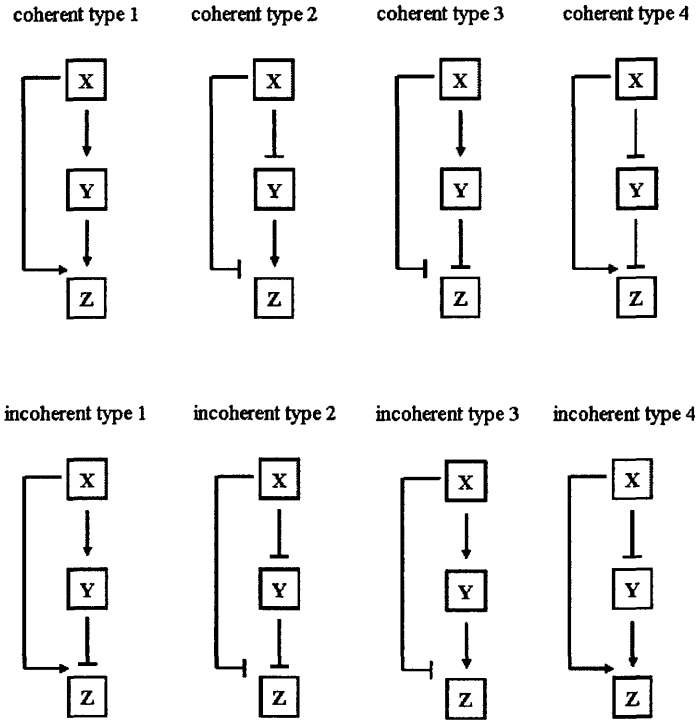

Figure 1-10: Structure of coherent and incoherent feedforward loop motifs.(Figure adapted from [4])

If the sign of the interaction between X and Z is the same as the overall sign of the chain of interactions between X, Y, and Z, then the motif is called 'coherent feedforward loop'. If, on the other hand, the signs are different, the motif is called 'incoherent feedforward loop'. The two categories of the feedforward loop motif are further specialized with respect to the type of the cis-regulatory logic at the Z pro-

moter: if both X and Y are necessary to express Z, then there is 'AND-gate logic at Z'; if either X or Y is enough to give rise to Z expression, then there is 'OR-gate logic at Z'. Feedforward motifs can be used to implement several functional characteristics into gene networks. For example, the type-1 feedforward motif with the 'AND-gate logic at Z' functions as a persistence detector: only for the persistent rather than transient inputs to X (or only for the persistent activation of the transcriptional activity of X), the motif would be functional.

1.5.4 Multi-input module

This motif consists of multiple regulators binding to a set of promoters to activate them. Therefore, this module could implement transcriptional bet-hedging strategy for cells. If a promoter can get activated with any one of a set of activators (which would be expressed at different levels in different environments), the cells would have a greater chance of having that promoter activated.

1.5.5 Single-input module

In this motif, a single transcriptional regulator is responsible from regulating the expression of several genes. This type of motif structure is frequently observed in metabolic pathways in which a set of genes take roles to metabolize a compound. For example [5], in the galactose utilization pathway of the budding yeast, the transcriptional regulator Gal4p binds to the promoters of several GAL pathway genes to activate their expression.

1.5.6 Regulator cascade

Regulator cascade motif is formed by a linear chain of regulator-promoter interactions, the first regulator binding to the promoter of a second regulator and that second regulator binding to the promoter of a third regulator, and so on. In the budding yeast genome, 188 regulator cascade structures have been found [3], each having between 3 to 10 regulators involved. Among the different network motifs observed

in the yeast genome, this motif is the one with the highest recurrence frequency. An important question to ask at this point would be what is so special about this motif so that it is utilized in gene networks much more frequently than other motifs. The answer to this question might lie in the functional characteristic of this network motif. Regulatory cascade motifs give the cellular networks the chance to separate the transcriptional events temporally. Only after the expression of the first regulator of the chain can the second regulator 'be in the game', providing cells with more control over the timing of transcriptional regulation.

1.6 Bistability in gene networks

Bistability in a gene network is set up by two gene expression states corresponding to the two activity levels of the network: ON (active) and OFF (inactive, basal). In order for a system to exhibit bistability [20], first, it has to have positive feedback or double negative feedback embedded into it. Second, the expression of the elements of the feedback loops should be nonlinearly regulated by their upstream regulatory proteins. For example, dimerization of an upstream regulatory protein could build such nonlinearity into the system. A bistable system will usually also exhibit hysteresis.

For a bistable gene network, hysteresis (Figure 1-11) is realized as history dependent gene expression level differences for isogenic cells exposed to the same environmental condition. The time-scale for hysteresis in a bistable system is set by the rates of stochastic transitions between the two states. If one waits long enough, no matter how strong the initial hysteresis is, the system (coming from different expression profiles corresponding to different histories) always ultimately reaches the same unique expression profile.

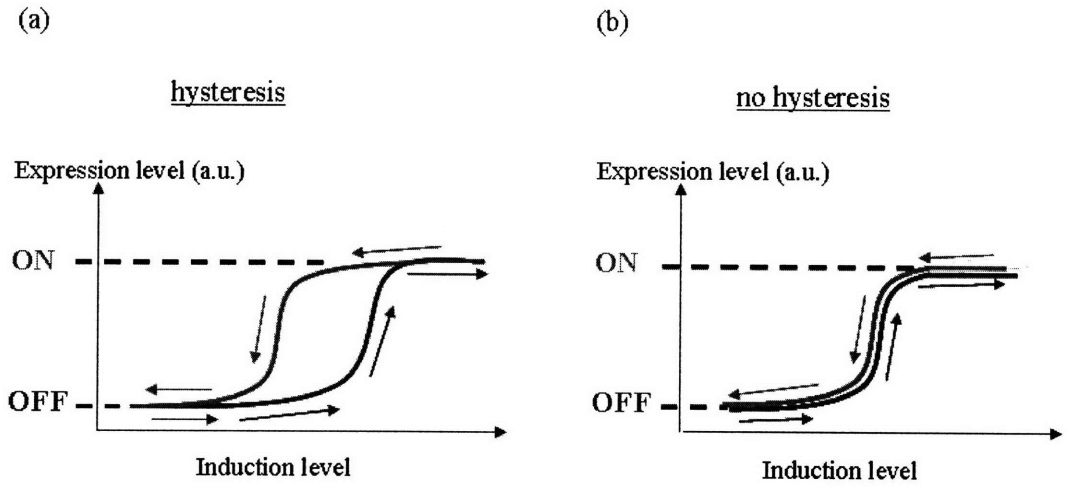

Figure 1-11: The presence and absence of hysteresis in a gene expression system. Red(blue) color indicates ON(OFF) history cells. a, Cells initially prepared in their ON and OFF states can give rise to different expression distributions for the same induction level. b, ON and OFF history cells show the same expression profiles at all times

Chapter 2

Enhancement of Cellular Memory by Reducing Stochastic Transitions

2.1 Summary

Upon induction of cell differentiation, distinct cell phenotypes are encoded by complex genetic networks [22–24]. These networks can prevent the reversion of established phenotypes even in the presence of significant fluctuations. In this chapter, using the yeast galactose-signalling network as a model system, the key parameters that determine the stability of cellular memory are explored. The galactose network contains multiple nested feedback loops. From the two positive feedback loops only the Gal3p-mediated loop is able to generate two stable expression states with a persistent memory of previous galactose consumption states. The parallel, Gal2p-mediated loop only increases the expression difference between the two states. A negative feedback through Gal80p reduces the strength of the core positive feedback. Despite this fact, a constitutive increase of the Gal80p concentration tunes the system from having destabilised memory to persistent memory. A model reveals that fluctuations are trapped more efficiently at elevated Gal80p levels. Indeed, the rate at which single cells randomly switch back-and-forth between expression states, was reduced. These observations provide quantitative understanding of stability and reversibility of cellular differentiation states.

2.2 Introduction

Complex gene and protein networks store cellular memory by creating two or multiple discrete, stable states of network activity [25–27]. Generation of bistability by simple feedback loops in synthetic circuits is well understood [19, 28–30]. However, naturally occurring networks, in particular in eukaryotic organisms, exhibit a complex organisation of multiple nested feedback loops making an analysis of system dynamics disproportionately more complicated [31, 32]. Such a network is exemplified by the galactose signalling pathway in the yeast *Saccharomyces cerevisiae*. Despite extensive data on its molecular interactions, an a priori prediction of its dynamical system behaviour is challenging [33, 34]. The galactose signal propagates through a four-stage signalling cascade. At the uppermost stage is the galactose transporter Gal2p, which imports extracellular galactose into the cell. Subsequently, intracellular galactose binds to and activates the cytoplasmic signal transducer Gal3p [35, 36]. At the third stage of this cascade, the activated Gal3p binds to and sequesters the inhibitor Gal80p to the cytoplasm depleting Gal80p from the nucleus [37]. The transcriptional activator Gal4p, which is constitutively bound to promoters of the GAL genes [38], is then released from the inhibitory action of Gal80p and activates expression of genes at the output of the cascade, including *GAL1*, *GAL2*, *GAL3*, and *GAL80*. Since an increase in Gal2p and Gal3p concentration results in enhanced transcriptional activity, these proteins close two positive feedback loops. The opposite holds for Gal80p that is part of a negative feedback loop. To read out the Gal4p activity in single yeast cells we monitored the expression of yellow fluorescent protein (YFP) driven by the *GAL1* promoter (Figure 2-1).

2.3 Bistability in the galactose network

Since the GAL regulatory network contains two positive feedback loops, this network has the potential to exhibit multistability. A convenient experimental way to probe for multistability is to subject the network to different initial conditions and explore

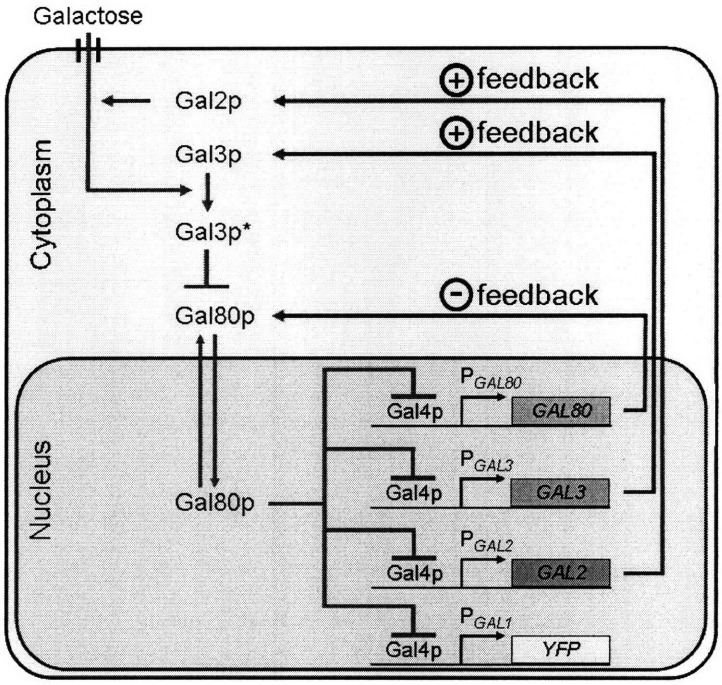

Figure 2-1: The galactose signalling pathway. Red arrows denote the four stage signalling cascade in which the external galactose signal controls the transcriptional activity of the GAL genes. The galactose bound state of Gal3p is denoted by Gal3p*. Pointed and blunt arrows reflect activation and inhibition, respectively. The double red arrows represent shuttling of Gal80p between the cytoplasm and the nucleus. The blue arrows denote feedback loops established by Gal2p, Gal3p, and Gal80p.

if the network gets locked in different stable expression states. Therefore, wild type cells were grown for 12 hours either in the absence of galactose ('raffinose history') or in the presence of 2 percent galactose ('galactose history') and subsequently grown for an additional 27 hours at a certain concentration of galactose. Raffinose was chosen as a neutral sugar since it neither activates nor represses the galactose regulated genes [39]. In order to maintain a constant concentration of galactose in the culture media, cells were grown at low densities so that galactose depletion was negligible (Appendix B). In Figure 2-2a expression histograms are shown for wild type cells for the two initial conditions and several galactose concentrations.

At low galactose concentrations cells exhibit a basal expression of YFP reflecting low Gal4p activity, whereas at larger galactose concentrations YFP expression is more than 100 fold higher, reflecting high Gal4p activity. Interestingly, the response to the two different initial conditions depends strongly on the galactose concentration. At low (< 0.012 percent) and high (> 0.35 percent) galactose concentrations the expression distributions after 27 hours do not depend on the history, and typically reach a steady-state after 6 hours. We classify this behaviour as history independent (absence of memory) since the system approaches the same unique expression distribution coming from different initial conditions. However for intermediate galactose concentrations the expression distributions obtained from the two different histories are significantly different and the system displays a memory of the initial galactose consumption state. We will classify this behaviour as history dependent (persistent memory) since cells get stably locked into two different expression states for periods much longer than the history independent system would need to reach steady state. Very similar behaviour is obtained when different promoters with Gal4p binding sites are used (Figure 2-6a) to drive YFP expression. Additionally, when the activity of two different GAL promoters are simultaneously monitored in single cells a strong correlation between promoter activities is observed (Figure 2-3). This implies that the transcriptional activity of the *GAL1* promoter and therefore YFP fluorescence, is a faithful reporter for Gal4p activity.

To pinpoint the network interaction responsible for the persistent memory we

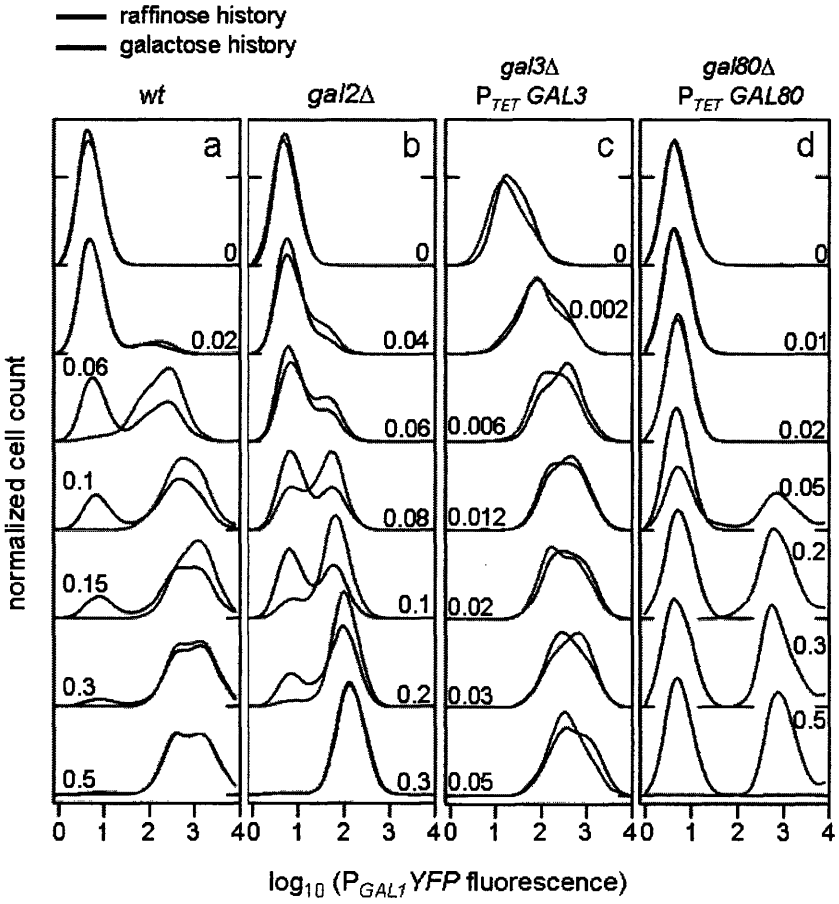

Figure 2-2: History dependent experiments reveal the regulatory interaction necessary for persistent memory. Expression distributions were obtained from at least 10000 cells and are plotted as a function of log(YFP fluorescence). Day-to-day variation in the expression histograms is typically less than 5 percent. Blue and red distributions denote cells that were initially grown for 12 hours without galactose and with 2 percent galactose, respectively. After this initial incubation cells were grown for an additional 27 hours in various concentrations of galactose as specified. a, Wild type strain MA0207. b, $\Delta gal2$ strain MA0215. The intensity of the high state is about 50 fold smaller than wild type. c, *GAL3* loop-knockout (MA0182). A doxycycline concentration of 0.05μg/ml was used. For this concentration the Gal3p expression corresponds to 80 percent of the Gal3p level observed in a wild-type strain induced with 0.5 percent galactose (Figure 2-4). d, *GAL80* loop-knockout (MA0188). A doxycycline concentration of 0.05μg/ml was used. For this concentration the Gal80p expression is very similar to the Gal80p level observed in a wild-type strain induced with 0.5 percent galactose (Figure 2-8).

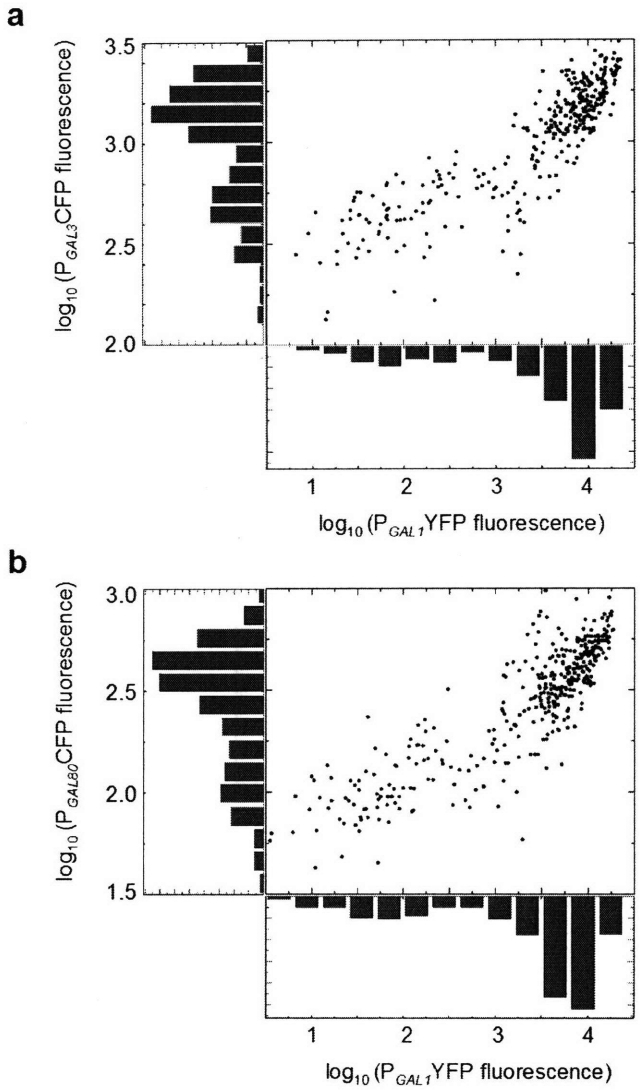

Figure 2-3: Single-cell two-color correlation experiments. a, Activities of *GAL1* and *GAL3* promoters driving YFP and CFP, respectively, were measured using fluorescence microscopy (strain MA0231). b, Activities of *GAL1* and *GAL80* promoters driving YFP and CFP, respectively, were measured using fluorescence microscopy (strain MA0242). 0.040 percent Galactose has been used to induce 'raffinose history' cells for 27 hours. Both histograms and the scatter plots indicate a strong correlation between the activities of the two promoters in single cells. We corrected for bleed-through between the CFP and YFP filters.

systematically interrupted the three feedback loops. Interestingly, the positive feedback loop mediated by Gal2p turned out not to be necessary for memory storage (Figure 2-2b). *gal2*Δ cells still display history dependent behaviour though the expression difference between the two states is reduced. In this case galactose is taken up by non-essential sugar transporters [40]. The non-essential nature of this feedback loop for memory storage contrasts prokaryotic metabolic networks where positive feedback through sugar transporters often defines the dynamics of the system [41].

To explore the role of the feedback loops through *GAL3* and *GAL80*, we constructed loop-knockouts in which the feedback loop was interrupted by replacing the endogenous, Gal4p dependent, promoter by an externally inducible, Gal4p independent, P_{TET} promoter. The capacity of storing the initial galactose consumption state was abolished by disrupting the *GAL3* loop. Irrespective of initial conditions the cell population approaches the same unique distribution (Figure 2-2c). The mean expression level increases gradually with increasing galactose concentration in contrast to the discrete transition observed for the wild type and *gal2*Δ cells (Figures 2-2a-b). This behaviour was observed for the entire *GAL3* expression range examined (from 5 percent to 300 percent with respect to wild-type Gal3p levels, Figure 2-4) and shows that feedback through Gal3p is necessary for memorizing the initial metabolic state in the wild type network.

A slow response of GAL genes has been observed in strains lacking a functional *GAL3* gene, several days after induction with galactose [42, 43]. The impact of this long-term adaptation depends strongly on which media or carbon source is used. In our media, having raffinose always present as a carbon source, long-term adaptation is not observed within 120 hours after adding galactose to the media and is therefore not relevant for the memory experiments (Figure 2-5).

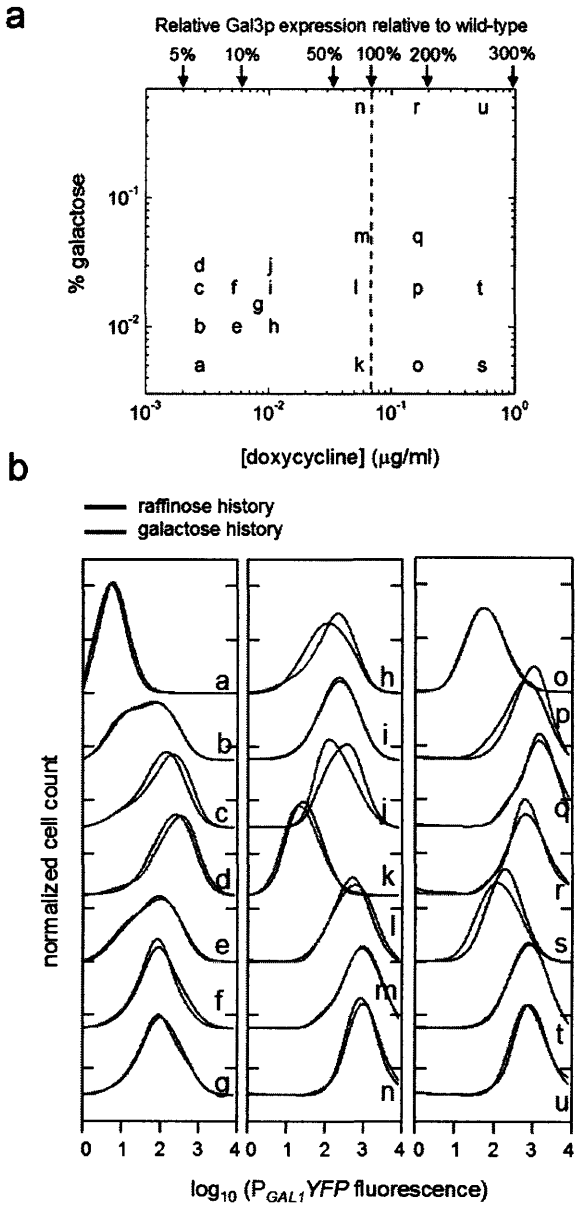

Figure 2-4: Loss of memory in the *GAL3* loop knock out (strain MA0182). a, Galactose and doxycyline concentrations are indicated for each experiments shown in b. The amount of Gal3p driven by the doxycycline inducible TET promoter is indicated relative to the wild-type expression (when induced with 0.5 percent galactose for 27 hours) as determined by fluorescence microscopy using Gal3p-CFP fusion proteins. The red dotted line represents the wild-type Gal3p level. b, MA0182 strain expression histograms (after 27 hr induction period) for a wide range of galactose and doxycyline concentrations.

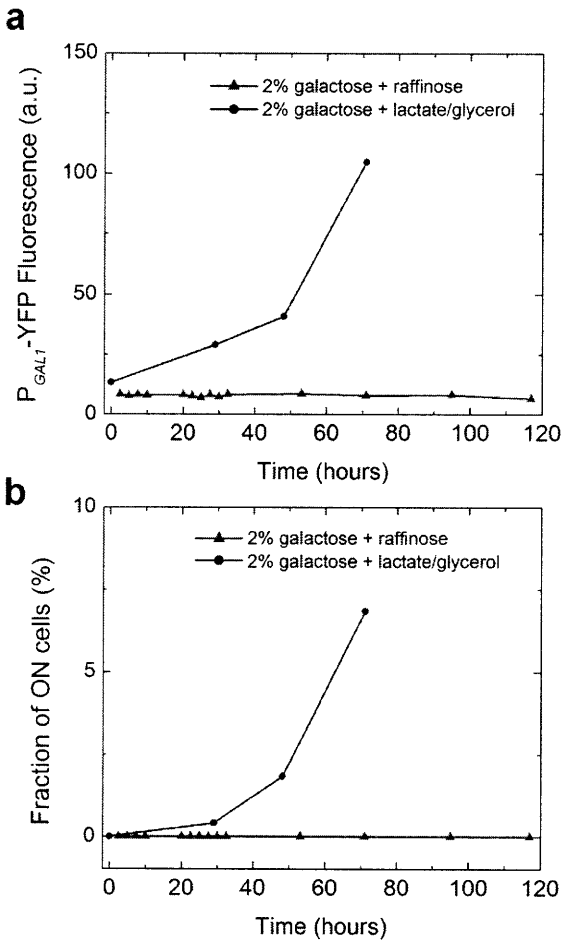

Figure 2-5: Occurrence of long-term adaptation in media with various carbon sources. a, Population averaged YFP levels in *gal3*Δ cells (strain MA0226) as a function of time for which the cells were induced with 2 percent galactose. b, Fraction of YFP expressing cells (strain MA0226) as a function of time. Long-term adaptation is not observed using synthetic media containing 2 percent raffinose and 2 percent galactose (red triangles). However when YEP is supplemented with 3 percent glycerol, 2 percent lactate, and 2 percent galactose (blue circles) long-term adaptation is observed as early as 30 hours after addition of 2 percent galactose to the media.

2.4 A phase diagram on the galactose vs. *GAL80* parameter space

When Gal80p was expressed constitutively at levels comparable to wild-type levels, the range of galactose concentration over which this system displays persistent memory significantly widened compared to wild type cells (Figure 2-2d). This is consistent with the earlier observation that the positive feedback through Gal3p is necessary for memory storage. In wild-type cells, the negative feedback through Gal80p effectively weakens the effect of the positive feedback through Gal3p and therefore decreases the persistent memory region in comparison to the *GAL80* loop knock out in which the negative feedback is abolished. For high concentration of galactose (> 0.1 percent) the expression distributions after 27 hours are indistinguishable from the initial conditions demonstrating extreme persistence of the initial expression states. For example, cells initially grown in the absence of galactose, still exhibit the same low Gal4p activity level even after 27 hours growth in 0.5 percent galactose. At this galactose concentration wild type cells would display maximum Gal4p activity (Figure 2-2a). In this regime the initial condition determines the future expression state, not the current concentration of galactose in the media.

This system was systematically explored over a broad range of *GAL80* expression (Figure 2-6b and Figure 2-8). In a large part of parameter space (galactose versus *GAL80* expression) the system displays persistent memory bordered by regions in which system memory is absent. In the latter case the system either approaches a state of low Gal4p activity (OFF) or high Gal4p activity (ON) independently of the history. Interestingly, at lower *GAL80* expression level a small region in parameter space was identified in which system behaviour displays destabilized memory. In this region the expression distribution displays two distinct peaks however the system response is not history dependent. We hypothesized that although the system has two stable states, fluctuations in gene expression of key regulators drive random transitions between the two states resulting in a destabilization of the memory. Indeed a significant cell-to-cell variation of gene expression levels was observed for several promoters in budding

yeast [11, 12].

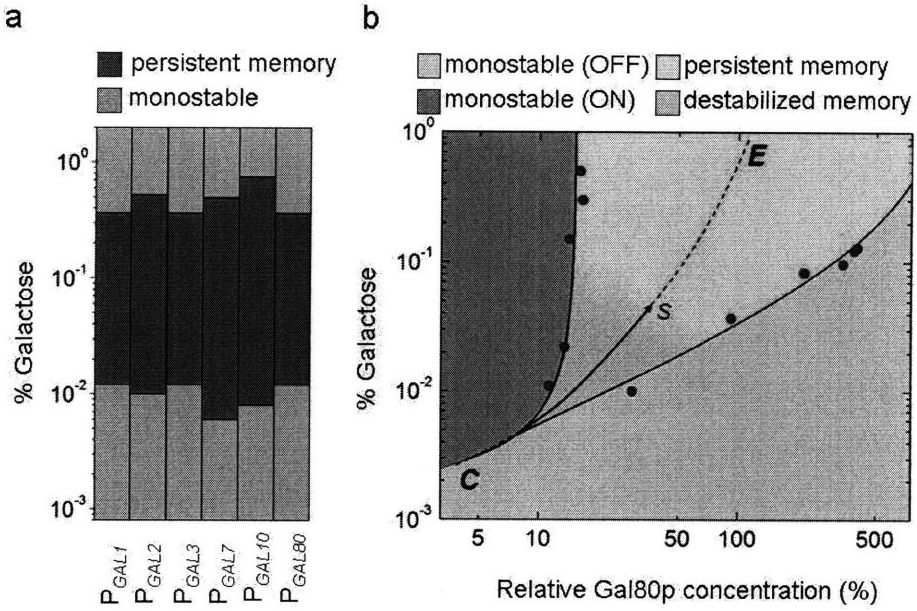

Figure 2-6: a, The range over which the system displays persistent memory is similar for different strains (from left to right column: MA0207, MA0208, MA0231, MA0212, MA0213, MA0242) in which the reporter gene is driven by different GAL promoters. b, System behaviour as a function of the control parameters: the external galactose concentration and the intracellular Gal80p concentration. *GAL80* expression is controlled by a doxycycline inducible promoter (MA0188) and is measured relative to wild-type Gal80p expression (induced by 0.5 percent galactose, Figure 2-8). Red circles indicate experimentally determined boundaries between different system behaviours and the black solid lines represent the theoretical prediction based on the regulatory network depicted in Figure 2-1. The critical point is defined as C. The coordinate s defined on the path from C to the endpoint E is used in Figure 2-7b to demonstrate the concept of energy barriers.

2.5 Predicting and experimentally testing the stability of cellular memory using the concept of energy landscapes

In order to reveal the effect of these fluctuations on the memory, the stability of the expression states was quantified using the concept of energy landscapes [44–

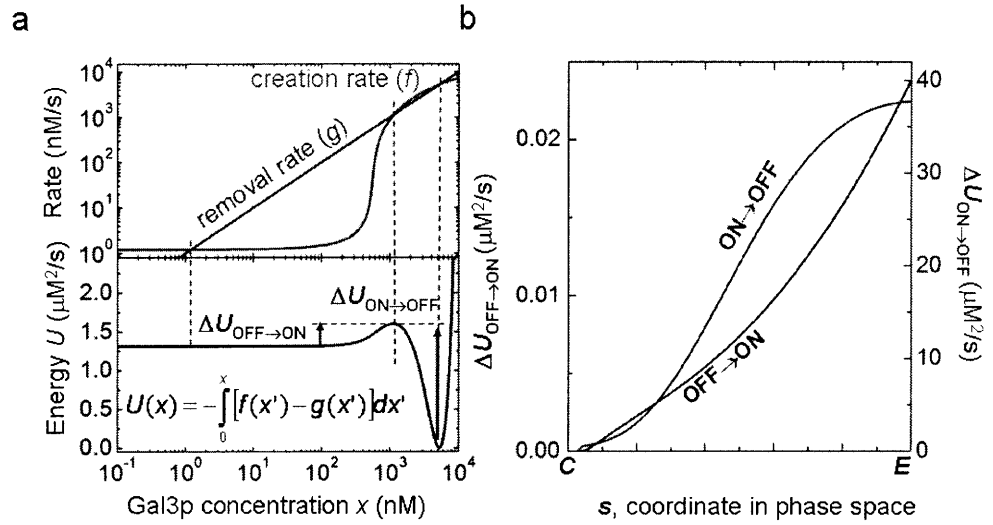

Figure 2-7: a, The energy landscape was calculated by integrating the difference between the creation f(x) and destruction rates g(x) of Gal3p with respect to the Gal3p concentration denoted by x. b, Magnitude of the energy barriers as a function of the coordinate s from C to E as defined in Figure 2-6b.

46]. The first-order differential equation describing the time evolution of the Gal3p concentration is analogous to the equation of motion of an overdamped particle in an energy landscape, where the particle position is analogous to the concentration of Gal3p. This analogy provides an intuitive representation for the stability of cellular expression states since minima in the energy landscape correspond to stable states that are separated by an energy barrier (Figure 2-7a). Fluctuations in gene expression are naturally introduced by analogy with the temperature experienced by the particle in the energy landscape. A particle trapped in an energy well at zero temperature will never escape, however at elevated temperatures the particle will display thermally activated transitions across the barrier. In this case the escape rate is proportional to $e^{-\Delta U/k_B T}$, where ΔU represents the energy barrier, k_B the Boltzmann constant, and T the absolute temperature [47]. The larger the energy barrier, the more efficiently the fluctuations are trapped in the vicinity of the stable states. The GAL system is characterized by two energy barriers: the barrier $\Delta U_{OFF \to ON}$ an OFF cell has to overcome to switch ON and the barrier $\Delta U_{ON \to OFF}$ for the opposite transition. The barrier height strongly depends on the system parameters. The barriers are high in

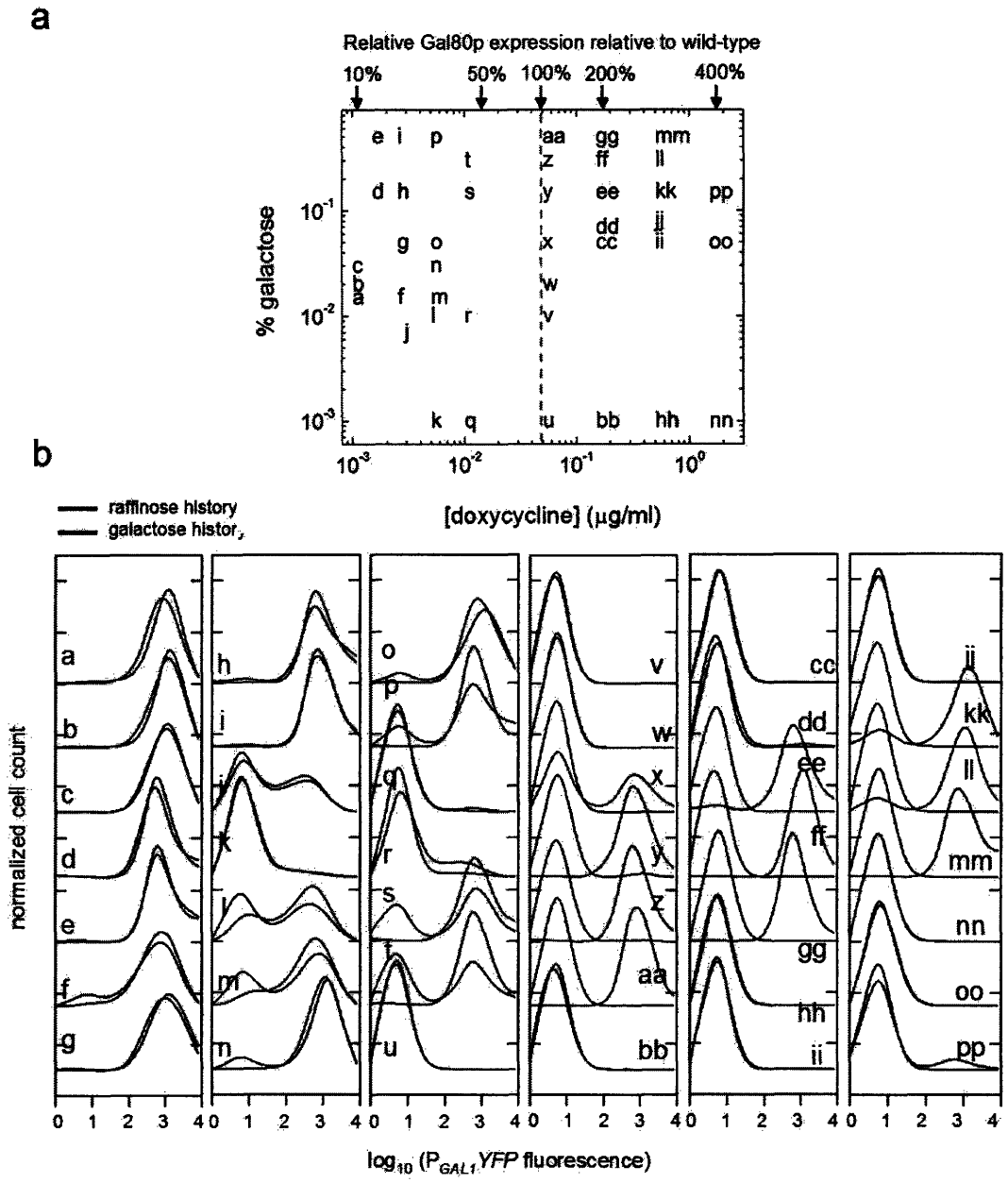

Figure 2-8: Behaviour of the *GAL80* loop knock out (strain MA0188). a, Galactose and doxycyline concentrations are indicated for each experiments shown in b. The amount of Gal80p driven by the doxycycline inducible TET promoter is indicated relative to the wild-type expression (when induced with 0.5 percent galactose for 27 hours) as determined by fluorescence microscopy using Gal80p-CFP fusion proteins. The red dotted line represents wild-type levels of Gal80p. b, MA0188 expression histograms (after 27 hr induction period).

the 'persistent memory' region and vanish as the critical point C is approached at lower *GAL80* expression level (Figure 2-6b and Figure 2-7b).

These predictions were verified by measuring the switching rates between the two states (Figure 2-9). For this purpose, we monitored the time evolution of the expression distribution as the system approached steady-state. An example of a bimodal distribution is shown in Figure 2-9a (top graph) obtained from the *GAL80* loop knock out operating in the 'destabilized memory' region (Figure 2-6b): some cells have high Gal4p activity while other cells display 100 times lower Gal4p activity. From this mixed population, two subpopulations were sorted that represent the two extremes of this distribution (Figure 2-9a, t = 0). Subsequently, these two subpopulations were separately grown in the same media as before sorting. The example in Figure 2-9a shows that both subpopulations relax back to the same bimodal distribution. These experiments therefore rule out static disorder and strongly suggest that cells switch back and forth between expression states. The switching rates strongly depend on the Gal80p level and the galactose concentration (Figure 2-9b). When the system operates in the 'destabilized memory' region close to the critical point C (Figure 2-6b) the system approaches steady-state in about 10 hours (Figure 2-9b, circles). However if the system operates closer to the boundary between the destabilized and persistent memory regions this process takes much longer (Figure 2-9b, triangles). At sufficiently high barriers, cells are almost irreversibly locked in one of the cellular states.

2.6 Escape rates as a function of energy barrier heights and the biological relevance of cellular memory

Figure 2-9c shows the experimental escape rates as a function of the calculated energy barrier. Escape rates drop precipitously as the height of the energy barrier increases, a correlation typical of noise induced transitions bounded by energy barriers. Interestingly, for the same barrier height, escape rates from the ON state are larger than

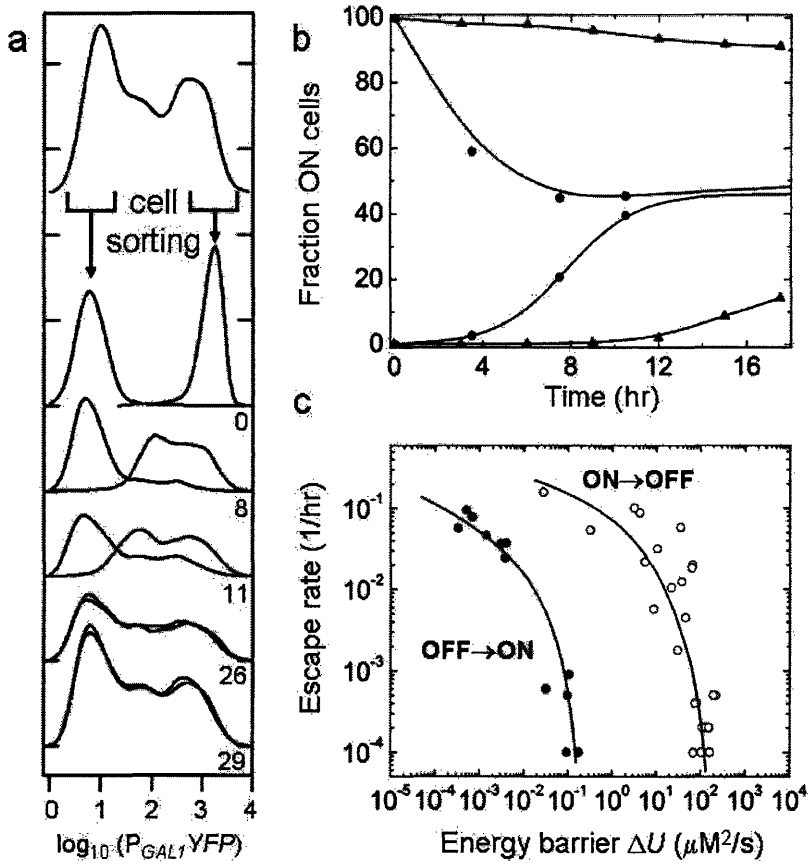

Figure 2-9: Stochastic switching dynamics between two stable expression states (strain MA0188). a, An initially bimodal population (0 percent galactose and 0μg/ml doxycycline, 27 hours of growth after raffinose history) was sorted by fluorescent activated cell sorting (FACS) into cells expressing low and high levels of YFP. At different times, indicated in hours after sorting, expression distributions were measured. b, The fraction of ON cells as a function of time after sorting in the absence of doxycycline and galactose (closed circles) and in the presence of 0.01μg/ml doxycycline and 0.040 percent galactose (closed triangles). Black and red solid lines (guides to the eye) reflect the sorted population that was initially fully OFF and ON respectively. The standard error on the fractions is smaller than 5 percent as determined by reproducing a representative histogram 10 times. c, Experimentally measured escape rates as a function of the calculated energy barriers (Section 2.8). Black lines and symbols indicate experimentally determined escape rates from the ON state whereas red lines and symbols denote escape rates from the OFF state. Energy barriers were calculated using experimentally determined parameters obtained from fitting the experimentally determined boundaries (Figure 2-6b, red circles) to the network model (Figure 2-6b, solid black lines, Section 2.8). The solid lines are guides to eye.

escape rates from the OFF state. Indeed the fluctuations in *GAL3* expression are larger in the ON state compared to the OFF state (Figure 2-10). However, given the complexity of this network and the presence of multiple noise sources [44, 45], it is unlikely that *GAL3* expression fluctuations can solely account for this difference in switching rates.

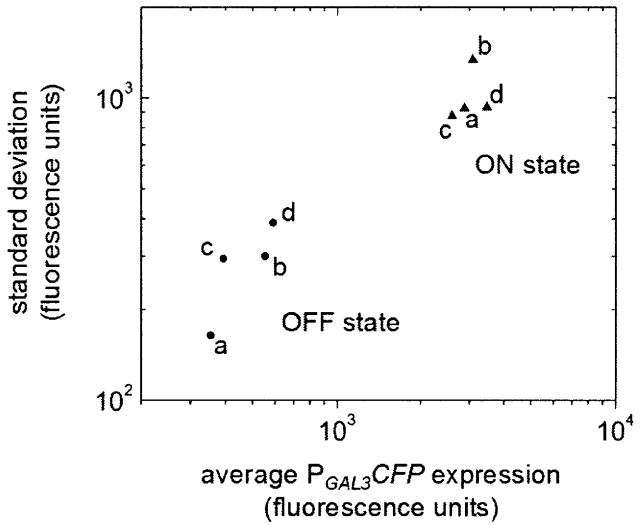

Figure 2-10: The role of Gal3p fluctuations in stochastic transitions. Raffinose history cells (strain MA0239) were induced with four different concentrations of galactose and doxycycline for 27 hours, each letter corresponding to a specific set of galactose and doxycycline: a 0 percent galactose, 0μg/ml doxycycline, b 0.004 percent galactose, 0.00185 μg/ml doxycycline, c 0.007 percent galactose, 0.0031μg/ml doxycycline, d 0.017 percent galactose, 0.00685μg/ml doxycycline. Fluorescence microscopy was used to determine CFP levels in individual cells. The resulting histograms were analyzed to determine the average P_{GAL3}CFP expression and the standard deviation (as a measure of *GAL3* fluctuations) for both OFF and ON state.

Next we explored if cellular memory could affect the growth rate of a population. Increased growth rate of the ON cells relative to the OFF cells was observed when the cells were grown in media in which galactose was the predominant carbon source (Figure 2-11). Since the fraction of ON cells in a population is strongly dependent on cellular memory, this indicates that memory might be an important concept in understanding fitness of a population in habitats in which carbon sources are fluctuating.

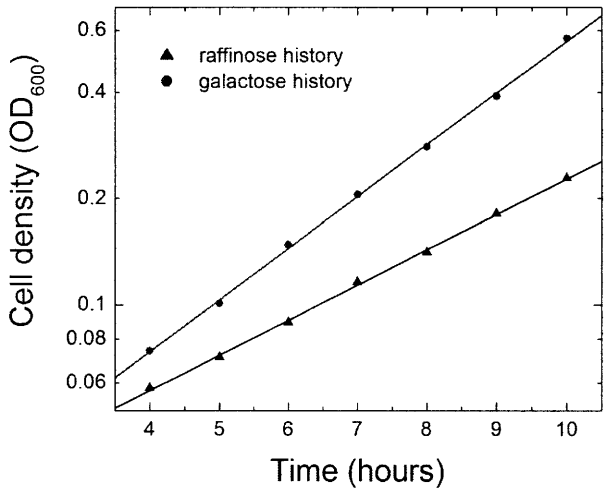

Figure 2-11: Growth rate is affected by history. Two populations of MA0188 cells ($gal80\Delta P_{TET}GAL80$) were prepared with a raffinose and galactose history in the persistent memory region (0.25 percent galactose, 0.15μg/ml doxycycline, 2 percent raffinose) for 27 hours. Subsequently these cells were washed and transferred to media lacking raffinose but having the same concentration of galactose and doxycycline (0.25 percent galactose, 0.15μg/ml doxycycline). After a lag phase of about 4 hours an exponential growth for both histories is observed, however, the galactose history cells divide at a rate that is about 1.5 fold larger than the raffinose history cells.

2.7 Discussion

The galactose signalling pathway of budding yeast has the potential to reliably store information on prior galactose exposures for hundreds of generations. A core positive feedback loop through *GAL3* is necessary for this cellular memory, whereas a negative feedback loop through *GAL80* competes with the positive *GAL3* loop and reduces the potential for memory storage. Consistently, when the negative feedback loop is opened and Gal80p levels are controlled constitutively, the memory persistence can be tuned from hours to months. A quantitative understanding of cellular differentiation would require a similar system-level approach as we employed for the galactose signalling pathway. Systematically opening the underlying feedback loops complemented by stochastic and stability analyses, provides a quantitative tool to identify network architectures responsible for the stability of differentiated cellular states.

2.8 Supplementary information

2.8.1 Modeling the galactose signaling pathway

The regulation of gene expression in the galactose signalling pathway was modeled by separating the fast reactions such as protein-protein and protein-galactose interactions from the relatively slow mRNA synthesis. Gal3p is a cytoplasmic protein [37], which gets activated upon galactose binding [35, 36]. The activated form of Gal3p will be denoted by gal•Gal3p. Gal80p shuttles between the nucleus and the cytoplasm [48]. Cytosolic Gal80p (Gal80p^C) binds to activated Gal3p (gal•Gal3p) whereas nuclear Gal80p, denoted by Gal80p^N, binds to and inhibits the transcriptional activator, Gal4p [37]. Gal4p is constitutively bound to the upstream activation sequences (UAS) of most of the GAL network genes [38]. UAS_{free} indicates the UAS concentration not bound to Gal80p, UAS•Gal80p denotes the DNA bound form of Gal80p. Gal80p is expected to exist as dimers since Gal80p dimerizes with high affinity [49]. The tri-molecular complex composed of Gal80p, Gal3p, and galactose is denoted by T. The fast reactions, assumed to be in equilibrium with respect to the relatively slow

transcriptional processes, are:

$$gal + Gal3p \longleftrightarrow^{\alpha} gal \bullet Gal3p$$
$$Gal80p^C + gal \bullet Gal3p \longleftrightarrow^{\beta} T$$
$$Gal80p^C \longleftrightarrow^{\gamma} Gal80p^N$$
$$UAS_{free} + Gal80p^N \longleftrightarrow^{\delta} UAS \bullet Gal80p^N$$

The corresponding equilibrium constants are defined by:

$$\alpha = \frac{[gal][Gal3p]}{[gal \bullet Gal3p]}$$
$$\beta = \frac{[gal \bullet Gal3p][Gal80p^C]}{[T]}$$
$$\gamma = \frac{[Gal80p^C]}{[Gal80p^N]}$$
$$\delta = \frac{[UAS_{free}][Gal80p^N]}{[UAS \bullet Gal80p^N]}$$

The total concentrations of the Gal3p, Gal80p, and UAS are given by the following mass-balance equations:

$$[Gal80p]_{Total} = [Gal80p^C] + [Gal80p^N] + [UAS \bullet Gal80p^N] + [T]$$
$$[Gal3p]_{Total} = [Gal3p] + [gal \bullet gal3p] + [T]$$
$$[UAS]_{Total} = [UAS_{free}] + [UAS \bullet Gal80p^N]$$

The last two sets of equations are now used to determine [UAS$_{free}$] for a given total concentration of Gal80p, Gal3p, UAS, and galactose. This involves solving the following equation:

$$Ax + \frac{Bx}{C+x} + \frac{Dx}{E+x} - F = 0$$

where

$$x \equiv [Gal80p^C]$$
$$A \equiv \frac{\gamma + 1}{\gamma}$$
$$B \equiv [Gal3p]_{Total}$$
$$C \equiv \beta(1 + \frac{\alpha}{[Gal]})$$
$$D \equiv [UAS]_{Total}$$
$$E \equiv \delta\gamma$$
$$F \equiv [Gal80p]_{Total}/F_o$$

In the experiments the total concentration of Gal80p is tuned by using a doxycycline inducible TET promoter. The relative concentration of Gal80p with respect to wild-type Gal80p levels (induced with 0.5 percent galactose) is determined by fluorescence microscopy using strains (MA0283 and MA0291) in which Gal80p is fused to CFP. Since these measurements provide only a relative measure, the parameter F_o was introduced as a scaling factor. The last equation above was solved using Matlab 6.0 (Mathworks, Inc.). Below we calculate the stability diagram for the Gal80p negative loop knock-out. The differential equation describing the dynamics of the core positive feedback by Gal3p is:

$$\frac{\partial[Gal3p]_{Total}}{\partial t} = k[UAS_{free}] - \Gamma[Gal3p]_{Total}$$

where k and Γ are the rate constants for creation and destruction of Gal3p, respectively. $[UAS_{free}]$ is determined by the solution of the above equation in x, $\tilde{x}$:

$$[UAS_{free}] = \frac{[UAS]_{Total}}{1 + \tilde{x}/E}$$

The final goal is to determine the number of solutions of the differential equation in the steady-state. Defining $\varepsilon \equiv DEk/\Gamma$ and $z \equiv [Gal3p]_{total}/\varepsilon$ and utilizing the

equation in x, the following polynominal in z holds:

$$D_3 z^3 + D_2 z^2 + D_1 z + D_0 = 0$$

where:

$$\begin{aligned} D_3 &= DE^2 - \varepsilon E - CDE \\ D_2 &= AE^2 - ACE - 2DE + EF + \varepsilon + CD - CF \\ D_1 &= -2AE + AC + D - F \\ D_0 &= A \end{aligned}$$

If this equation has one real root the system is monostable, if it has 3 real roots it is bistable. The boundaries between the monostable and bistable regime are found by realizing that at the boundary the polynomial in z has exactly two real solutions and therefore it can be written as:

$$z^3 + \frac{D_2}{D_3} z^2 + \frac{D_1}{D_3} z + \frac{D_0}{D_3} = (z-a)^2 (z-b) = 0$$

This results in the following conditions:

$$\begin{aligned} \frac{D_2}{D_3} &= -(2a+b) \\ \frac{D_1}{D_3} &= a^2 + 2ab \\ \frac{D_0}{D_3} &= -a^2 b \end{aligned}$$

Using the first two conditions results in two solutions for a:

$$\begin{aligned} a^+ &= \frac{1}{3}\frac{D_2}{D_3}\left[\sqrt{1 - \frac{3D_3 D_1}{D_2^2}} - 1\right] \\ a^- &= -\frac{1}{3}\frac{D_2}{D_3}\left[\sqrt{1 - \frac{3D_3 D_1}{D_2^2}} + 1\right] \end{aligned}$$

In principle the last two sets of equations can now be solved implicitly and the boundaries are obtained. However an accurate explicit form of the boundaries is found by approximating the last two equations with a second order Taylor expansion with respect to $3D_3D_1/{D_2}^2$. The advantage of the explicit method is that the experimental boundaries can be fitted directly without numerically solving the last two sets of equations above. Figure 2-12 demonstrates the accuracy of the approximate method.

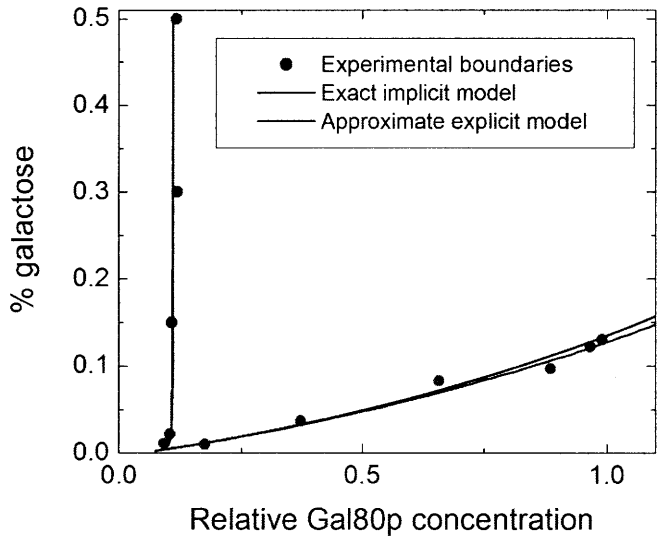

Figure 2-12: Phase boundaries determined by the approximate explicit method and the exact implicit method. For the left boundary both methods yield indistinguishable results. For large relative Gal80p concentrations the approximate result deviates slightly with respect to the exact result.

Substituting a^+ into $D_0/D_3 = -a^2b$ gives the condition for one of the phase boundaries:

$$-64D_1^2D_2^8 - 32D_1^3D_2^6D_3 + 108D_1^4D_2^4D_3^2 + 108D_1^5D_3^3D_2^2 + 27D_1^6D_3^4 + 256D_0D_2^9 = 0$$

Since terms with high orders of D_2 dominate, an adequate approximation is:

$$D_1^2 = 4D_0D_2$$

Similarly, by substituting a^- into $D_0/D_3 = -a^2b$ the other boundary is found. The condition for the second boundary is:

$$2D_2^2 = 9D_3D_1$$

The last two relations were fitted to the experimental data using α, F_o, and k/Γ as fit parameters. The equilibrium constants $\beta = 0.06$ nM and $\delta = 0.05$ nM were obtained from literature [49, 50]. The shuttling constant $\gamma \approx 1$ was estimated from localization studies of GFP-Gal80p [37, 48]. The total concentration of UAS was assumed to be 50 nM based on the total number of GAL4 binding sites in the yeast genome [51] and an approximate nuclear volume of $1\mu m^3$. The experimentally obtained boundaries (red circles, Figure 2-6b) were best fitted using the parameters $\alpha = 1$ percent, $F_o = 520\mu$M, and $k/\Gamma = 470$. The theoretical boundaries are denoted by the solid black lines in Figure 2-6b. These parameters were subsequently used to calculate the energy landscapes and energy barriers as defined in Figure 2-7a :

$$U \equiv -\int_0^{[Gal3p]_T} (k[UAS_{free}] - \Gamma[Gal3p]_{Total}) d[Gal3p]_{Total}$$

Note that $[UAS_{free}]$ is a function of $[Gal3p]_{Total}$. Therefore, this integral has to be evaluated numerically.

2.8.2 Determining experimental escape rates

The dynamics of noise induced transitions between the two stable states of a bistable system can be studied by a three parameter model composed of the fraction of the cells in the OFF or ON states and the rates specifying the likelihood of the transitions between those two states. In what follows, the fraction of OFF-cells is represented by a_0, and that of ON-cells by $a_1 = 1 - a_0$. The forward and backward switching rates are denoted by α_+ and α_-, respectively. The growth rates associated with the OFF and ON cells are assumed to be equal which is consistent with the experimentally obtained growth rates. In this analysis the stability of the fluorescent reporter YFP

is ignored. Fluorescent reporters are typically very stable. However cells grow at a rapid rate and therefore dilute out the intracellular YFP at a rate that equals the cell growth rate. Since the exponential dilution rate due to cell growth (0.5 hour^{-1}) is much faster than the typical stochastic switching rates ($\ll$ 0.1 hour^{-1}, Figure 2-9c), the switching rates are adequately approximated. The switching dynamics between the two states is described by the following differential equations:

$$\frac{da_0}{dt} = \alpha_- a_1 - \alpha_+ a_0$$
$$\frac{da_1}{dt} = \alpha_+ a_0 - \alpha_- a_1$$

Since $a_0 + a_1 = 1$, the solution to these differential equations is:

$$a_1(t) = a_1(0) + (\frac{\alpha_+}{\alpha_+ + \alpha_-} - a_1(0))(1 - e^{-(\alpha_+ + \alpha_-)t})$$

where $a_1(0)$ is the initial fraction of ON cells. Cells pre-grown with and without galactose will have different initial fractions defined by $a_1(0)_G$ and $a_1(0)_R$, respectively. Subtracting $a_1(t)_R$ from $a_1(t)_G$ leads to:

$$\alpha_+ + \alpha_- = \frac{ln(\frac{a_1(0)_G - a_1(0)_R}{a_1(t)_G - a_1(t)_R})}{t}$$

Substituting this equation in $a_1(t)_R$, $\alpha_+/(\alpha_+ + \alpha_-)$ is found:

$$\frac{\alpha_+}{\alpha_+ + \alpha_-} = \frac{a_1(t)_R - a_1(0)_R + a_1(0)_R(1 - \frac{a_1(t)_G - a_1(t)_R}{a_1(0)_G - a_1(0)_R})}{1 - \frac{a_1(t)_G - a_1(t)_R}{a_1(0)_G - a_1(0)_R}}$$

Since the initial fractions ($a_1(0)_G$, $a_1(0)_R$) and fractions at a specific time ($a_1(t)_G$, $a_1(t)_R$) are experimentally known, the escape rates α_+ and α_- for the two stable states can be estimated.

2.8.3 Doxycycline inducible P_{TET} system

To externally control the Gal3p and Gal80p levels in strains MA0182 and MA0188 respectively, the *GAL3* and *GAL80* genes were placed under the control of the P_{TET} promoter. In the presence of doxycycline, this promoter is activated by the transcriptional activator rtTA, comprised of the VP16 activator and the TetR(S2) DNA binding domains. rtTA was chosen to control the Gal3p and Gal80p levels, since it is a transcriptional activator with a broad regulatory range [52].

2.9 Methods

2.9.1 Plasmid and strain constructions

*Kpn*I-promoter-*Bam*HI, *Bam*HI-YFP-*Eco*RI fragments were cloned into pRS402 backbone upstream of CYC1 transcriptional terminator (for P_{GAL4} promoter, *Bgl*II was used instead of *Bam*HI). The P_{GAL1}, P_{GAL2}, P_{GAL3}, P_{GAL4}, P_{GAL6}, P_{GAL7}, P_{GAL10}, P_{GAL80}, and P_{MYO2} promoter sequences correspond to 669, 960, 830, 633, 590, 724, 667, 632, and 677 base pair regions upstream of the start codon of the respective genes. The *GAL3* and *GAL80* genes were cloned downstream of *Kpn*I-P_{TETO2}-*Bam*HI as *Bam*HI-*Eco*RI fragments into pRS306. P_{GAL3}-*GAL3*, P_{GAL80}-*GAL80*, P_{TETO2}-*GAL3*, P_{TETO2}-*GAL80* missing the stop codon were cloned upstream of *Bam*HI-CFP-*Eco*RI as *Kpn*I-*Bam*HI fragments in pRS306. All strains were derived from W303. PCR and Southern blot were used to verify the integrations. Complete descriptions of the strains used in this study can be found in Appendix A.

2.9.2 Growth conditions and media

Cultures were grown in synthetic dropout media with the appropriate aminoacid supplement and 2 percent raffinose as a carbon source. Media used for 'galactose history' experiments were supplemented with 2 percent galactose, while media for 'raffinose history' experiments contained raffinose as the sole carbon source. Overnight grown cells were diluted to an OD_{600} value such that after 27 hours growth, the galactose

concentration in the media would not change by more than 10 percent (section 2.9.4). Cells were grown at 30°C. After the induction period of 27 hours, the expression distributions were determined by flow cytometer (FACScan, Becton-Dickinson). For cell sorting, the bimodally distributed cells were sorted into OFF and ON cells as indicated in Figure 2-9a.

2.9.3 Gal3p and Gal80p levels in inducible strains relative to wild-type levels

To determine how Gal3p and Gal80p levels in strain MA0182 and MA0188 respectively, compare to native levels of Gal3p and Gal80p in wild type cells, strains were constructed in which the C terminus of Gal3p and Gal80p was fused to CFP. Using fluorescence microscopy the levels of doxycycline-induced *GAL3* and *GAL80* expression relative to the wild-type expression were determined. We find that addition of 0.07μg/ml and 0.05μg/ml doxycycline results in expression of wild-type levels (wild-type cells fully induced with 0.5 percent galactose) of Gal3p and Gal80p, respectively (Figure 2-4 , Figure 2-8).

2.9.4 Determination of galactose consumption rate

To determine the galactose consumption rate aliquots from cultures were filtered and the galactose concentration of the cell-free media was analysed as follows. β-galactose-dehydreogenase was used to oxidise galactose in the presence of 2.5 mM NAD dissolved in a buffer containing 50 mM imidazole, 5 mM $MgCl_2$ pH=7.0 [53]. Conversion of NAD into NADH was followed spectrophotometrically at 340 nm (Appendix B).

Chapter 3

Phenotypic Bet-Hedging as a Survival Strategy in Yeast Populations

3.1 Summary

A classic problem in evolutionary and population biology is to understand how a population optimizes its fitness in fluctuating environments. Rather than maintaining a phenotypically homogenous population, it has been suggested that a 'bet-hedging' strategy might be more beneficial in uncertain environments. Following this strategy, a population consisting of a variety of phenotypes enhances its fitness by ensuring that, at any given time, some of its members are prepared for an unforeseen environmental fluctuation. We experimentally test this hypothesis in vivo using a re-engineered yeast strain that randomly transitions between two phenotypes. Each phenotype is designed to confer a growth advantage over the other phenotype in a certain environment. We show that, in order to optimize population growth, cells have to match their inter-phenotype switching rate to the frequency of environmental changes. Our data suggest that when transition rates are correctly tuned, random bet-hedging constitutes a simple, yet effective, survival strategy to cope with fluctu-

ating environments without the need to actively sense environmental conditions.

3.2 Introduction

Unlike controlled laboratory environments, cells in the wild have to face and surmount the challenges raised by random fluctuations in extracellular conditions [54–57]. Signal transduction pathways allow cells to actively sense and respond to particular environmental changes. However, given the myriad of environmental conditions and the limited number of signal transduction pathways, it is likely that cells use complementary, more passive methods to anticipate changes in the environment. Without the need to sense the environment, cells can 'blindly' anticipate and survive environmental changes by randomly switching between multiple phenotypes, each fit to a particular environment. Following this bet-hedging strategy [56, 58, 59], a population consisting of a variety of phenotypes enhances its fitness by ensuring that, at any given time, at least some of its members are prepared for an unforeseen environmental fluctuation.

This random bet-hedging is naturally employed by many microbial systems in order to cope with environmental uncertainty [16]. One prominent example is known as phase variation [60], which defines two distinct phenotypes between which cells reversibly transition. Phase variation in the expression of surface pili of *Escherichia coli* during infection of the urinary tract allows single cells to have two phenotypes: a free-floating phenotype and a phenotype that adheres to the urinary tract surface [61, 62]. Similarly, the methylation state of the agn43 promoter of *E. coli* defines two phenotypes: a individual planktonic life-style or community life in the biofilm [63]. More examples include the transition to competence in *Bacillus subtilis* [64, 65], bistability in metabolic networks [41,66], and antibiotic tolerance in many prokaryotes [67]. Transitions between the two distinct phenotypes are thought to be random and rare, meaning that the transition frequency is much smaller than the cell division rate. Therefore an important question which arises in this context is how the inter-phenotype transition rates related to the fitness of the population.

Several recent theoretical models, developed to address this topic, have proposed that an optimum bet-hedging strategy is achieved when the rate of stochastic transitions between phenotypic states is tuned to match the rate of fluctuations in the environment [14–16, 68, 69]. Here we experimentally test this hypothesis in vivo using a re-engineered version of the galactose utilization network of the budding yeast *Saccharomyces cerevisiae*. We use this particular re-engineered network because it allows a more quantitative control over both the inter-phenotype switching rates and phenotypic growth rates than would be possible using one of the natural systems where bet-hedging is thought to play an important role. Stochastic fluctuations in gene expression [5, 8–10, 70] in the re-engineered network allow a single cell to randomly switch back-and-forth between two distinct phenotypes. Each phenotype is designed to confer a growth advantage over the other phenotype in a certain environment. We experimentally demonstrate that in order to optimize population growth, cells have to match the inter-phenotype switching rate to the frequency of environmental changes. Our experiments demonstrate that random bet-hedging can constitute a simple, yet effective, survival strategy to cope with fluctuating environments without the need to actively track environmental conditions.

3.3 A discrete switching model for random bet-hedging

In general, cells are faced with a continuous spectrum of dynamically changing environmental conditions, including fluctuations in temperature, pH, and concentrations of nutrients and toxins. To cope with these diverse environmental conditions, cells could potentially explore a multitude of phenotypes. Here we focus on a simpler, discrete system with two environments and two phenotypic states (Figure 3-1a) that captures the important properties of bet-hedging strategies [14, 16, 69].

If the stochastic switching between the two phenotypic states is much faster than the switching between the two environmental states, a high level of phenotypic diver-

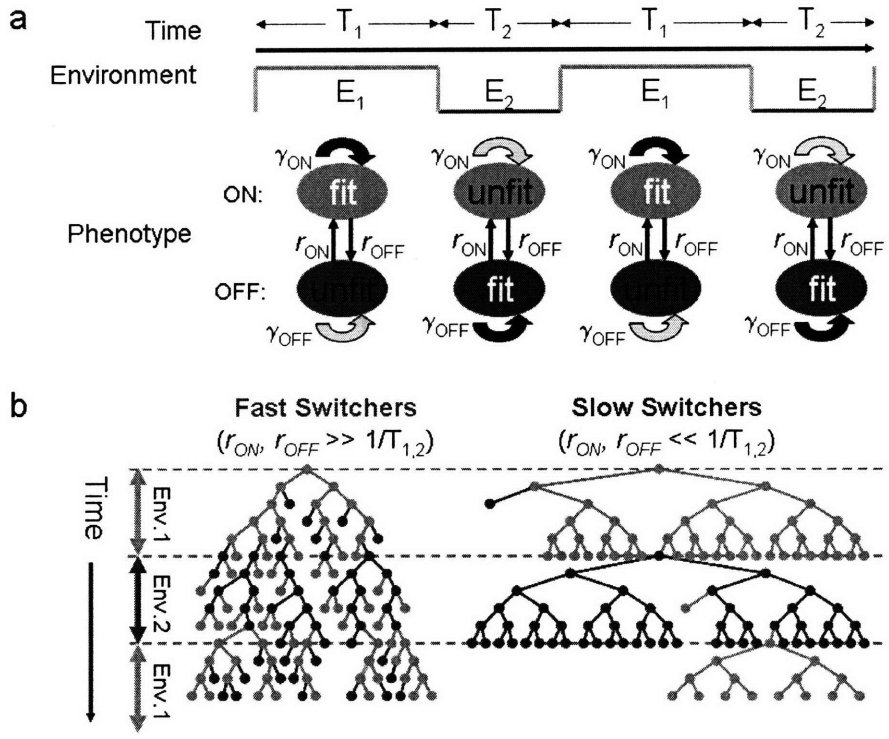

Figure 3-1: a, Two states (phenotypes) exist for each cell, ON (orange) and OFF (green). Cells randomly switch between the two states with frequencies r_{ON} and r_{OFF}. The first environment (E1) has no uracil, while the second (E2) has both 5-FOA and uracil. A cell is either fit or unfit to its environment depending on the specific phenotype it displays. For example in E1, on-state cells are fit with a growth rate, γ_{ON}, but the unfit off-state cells proliferate with a smaller growth rate, γ_{OFF}. b, Cellular lineage for fast and slow switchers. Single cells with different switching frequencies (fast and slow) grow in alternating environments. Orange and green colors represent the first and second environments, respectively. Color change in the cellular lineage diagram corresponds to the change in phenotypic expression for a particular cell. If a cell finds itself in the unfit state after a switching event, it ceases to proliferate. In the case of slow switchers, reduced cell-to-cell variability in each environment is depicted by a more dominant use of a single color.

sity is expected (Figure 3-1b, left panel). However, if the phenotypic switching rate is much slower than the environmental switching rate, the population is more homogenous at any given point in time (Figure 3-1b, right panel). In a constant environment, the slow switchers maintain larger growth rates than fast switchers due to a greater fraction of cells in the fit state. However, slow switchers suffer greater losses and take more time to recover whenever an environmental transition occurs. In other words, a population must balance the rate at which cells transition between the phenotypic states with the frequency of environmental changes to maintain a reasonable number of cells in the unfit phenotype. This suggests that in order to optimally benefit from the diversity, cells should tune their switching rates to match the frequency of environmental changes [14–16, 68, 69]

3.4 Re-engineering the galactose utilization network

To experimentally test this hypothesis we implemented this discrete switching system (Figure 3-1a) *in vivo*. Bistable gene networks [31, 71, 72] provide promising experimental systems to implement discrete phenotypic states. The stochastic nature of gene expression drives rare transitions between these states [9, 10, 44, 46, 70]. We utilize the bistable galactose utilization pathway of the budding yeast *Saccharomyces cerevisiae* [5, 73], because the switching rates between the two phenotypic states can be tuned experimentally giving us precise control over the system [5].

The wild type network was re-engineered as shown in Figure 3-2a. The activity of the pathway was read out at the single cell level using yellow fluorescent protein (YFP) under the control of the *GAL1* promoter.

For certain extracellular galactose concentrations, cells in an isogenic population display either a basal pathway activity (OFF) or an approximately 100-fold up-regulated activity (ON). Figure 3-2b and Figure 3-2c illustrate the stochastic transitions between these phenotypic states in single cells. Here a single progenitor cell

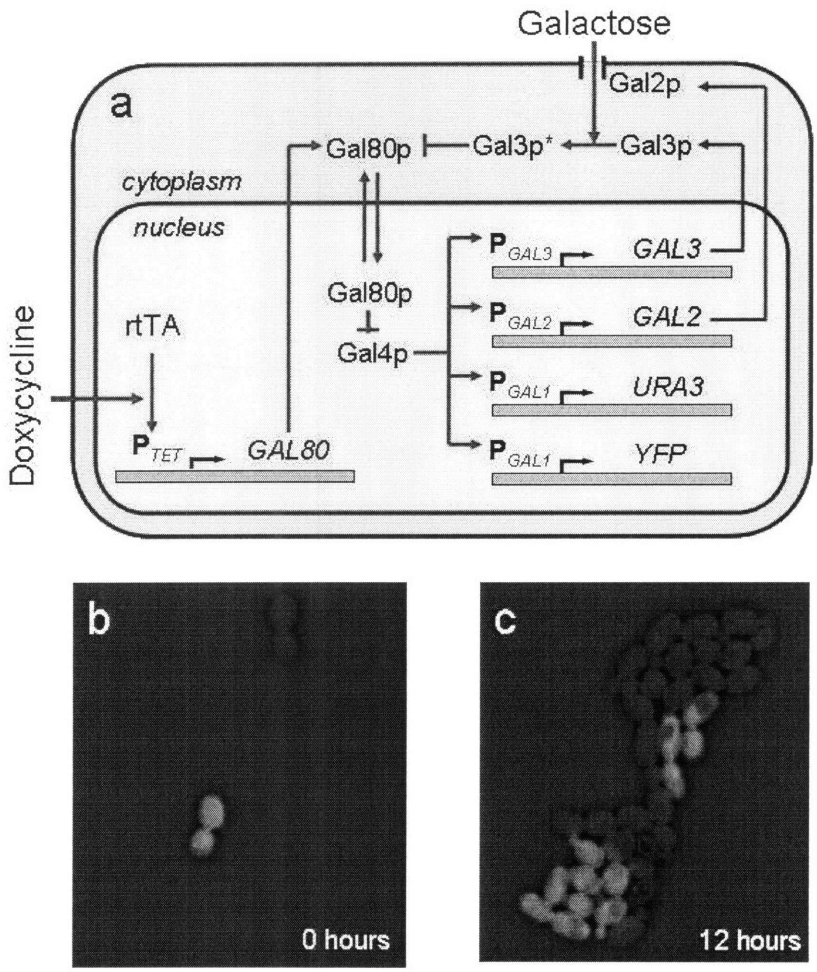

Figure 3-2: a, The galactose-signaling pathway. The activity of the galactose pathway is read out by using YFP driven by the *GAL1* promoter. Similarly, the endogenous *URA3* expression is also under the control of the *GAL1* promoter, coupling the synthesis of the Ura3 proteins to the activity of the GAL pathway. In environment E1, ON cells will synthesize uracil and thrive, while in environment E2 the *URA3* gene product is converted to a toxic intermediate in the presence of 5-FOA. By changing the extracellular galactose and doxycycline concentrations, the transition rates between the ON and OFF states can be altered, providing us with the fast and slow switchers. b, A single cell initially in the OFF state gives rise to both ON (green) and OFF (black) cells due to stochastic transitions between states. c, Similarly an ON cell (green) gives rise to a mixed population.

initially in either the ON or the OFF state gives rise to a diverse population of both ON and OFF cells due to stochastic transitions between the two phenotypes.

In our system, *GAL2* and *GAL3* remain under the control of their endogenous promoters, while the expression of *GAL80* is placed under the control of the *TET* promoter, which is induced using doxycycline. Because of this modification, this strain has externally tunable stochastic transition rates, r_{ON} and r_{OFF} (Figure 3-1a), between the two phenotypic expression states. The rates are controlled by changing the extracellular concentration of galactose and doxycycline, with low (high) amounts of galactose and doxycyline leading to high (low) transition rates [5]. This feature allows us to create populations of cells with both fast and slow switching rates.

To provide the two phenotypic states with two distinct growth rates we placed the endogenous *URA3* expression, necessary for uracil biosynthesis, under the sole control of the *GAL1* promoter. In this scenario, the two environments (E1 and E2) affect the growth rates of the two phenotypes in an antagonistic manner. Environment E1 lacks uracil and therefore favors the growth of ON cells to the disadvantage of the OFF cells. In contrast, E2 contains both uracil and the small molecule 5-Fluoroorotic acid (5-FOA), which is converted into the toxic intermediate 5-Flurouracil in the presence of Ura3 protein, conferring a growth advantage to OFF cells [74]. Therefore, this experimental system provides us with a quantitative control over both the inter-phenotype switching rates and the growth rates of both phenotypes in the two different environments.

3.5 Characterization of inter-phenotypic switching rates and antagonistic selection

We first create two distinct populations of cells: one with fast and one with slow phenotypic transition rates, r_{ON} and r_{OFF}, by externally controlling the concentration of galactose and doxycycline present in the extracellular medium. Figures 3-3a-b experimentally demonstrate the behavior of these two populations in non-selective

media.

Due to stochastic transitions between the two phenotypes, the t=0 expression distributions corresponding to fast (0.004 percent galactose, 0.00282μg/ml doxycycline) or slow (0.03 percent galactose, 0.0135μg/ml doxycycline) switching cells change at the end of the 24 hours growth period. By measuring the fraction of cells in each state at several points in time (Figure 3-4), we are able to estimate that the transition rates are roughly ten times greater for the fast switching cells, ($r_{ON} \approx 0.047\text{hr}^{-1}$, $r_{OFF} \approx 0.035\text{hr}^{-1}$), than for the slow switching population, ($r_{ON} \approx 0.004\text{hr}^{-1}$, $r_{OFF} \approx 0.007\text{hr}^{-1}$).

Figures 3-3c-d experimentally demonstrate the selection pressures imposed by each environment on the two phenotypes, ON and OFF. In the absence of any selection, both OFF and ON cells grow at very similar rates. In this case, for both the fast and slow switchers we observe steady-state bimodal distributions of network activity with roughly equal numbers of cells in either phenotypic state (Figs 3-3c-d, top panels). However, when cells are grown for 4 days in either environment E1 or E2, the population significantly enriches in ON or OFF cells, respectively (Figs 3-3c-d, bottom panels). The fast switchers display a more diverse distribution of expression values compared to the slow switchers, because the fast switchers transition more frequently to the unfit phenotype.

3.6 Growth rate dynamics of fast and slow switchers

Next, we explored how this increased diversity impacts the growth rate of the population. To measure growth rates over long periods we utilized turbidostats [75] enabling accurate and automatic on-line measurements of population growth rates (Figs 3-5a-b).

Cells are maintained at a constant optical density in liquid culture by continuous measurement of the culture's relative absorbtion coefficient (Figure 3-5b, upper panel

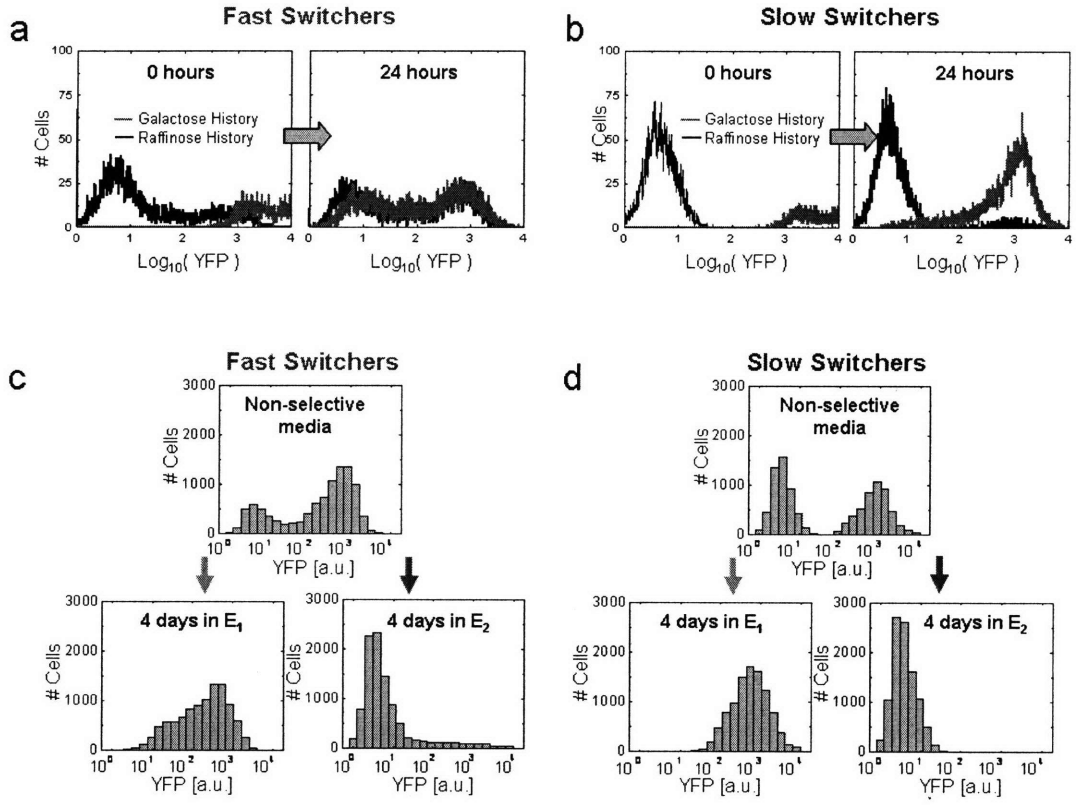

Figure 3-3: a, YFP fluorescence distribution of fast switching cells that have been grown in non-selective media to obtain the t = 0 distributions (Section 3.10). After an additional 24 hours in non-selective media (with 0.004 percent galactose and 0.00282μg/ml doxycycline corresponding to fast switchers), the distributions show very little history dependence indicating that cells are switching much faster than 1/24 hours^{-1}. b, YFP fluorescence distribution of slow switching cells that have been grown in non-selective media to obtain the t = 0 distributions (Section 3.10). After an additional 24 hours in non-selective media (with 0.03 percent galactose and 0.0135μg/ml doxycycline corresponding to slow switchers), the distributions show a large history dependence indicating that the switching rates are much slower. c, Fast switching cells grown in non-selective media display a bimodal distribution. When cells are grown in E1 (E2), the interaction of the *URA3* gene with the environment causes ON (OFF) cells to proliferate. d, Similar selection is observed for slow switching cells, however fewer unfit cells are observed compared to the fast switchers.

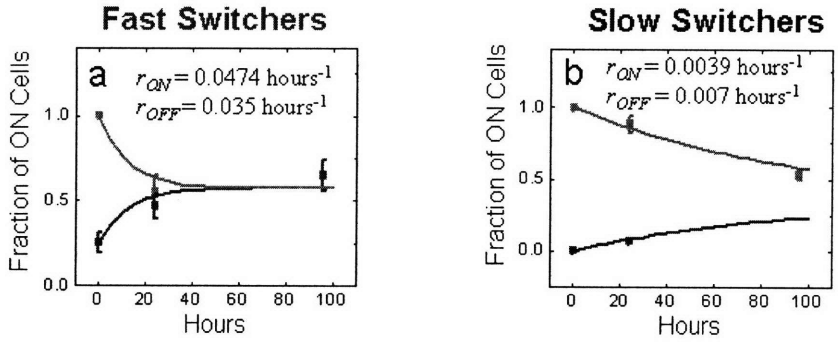

Figure 3-4: Determination of switching rates for fast and slow switchers. The fraction of ON cells as a function of time in non-selective media for the galactose (orange) and raffinose (green) history cells. To estimate the fraction at 96 hours the steady state distribution of cells (taken from Figs 3-3c-d) were determined. By fitting the data with the function: $f_{ON}(t) = \frac{r_{ON}}{r_{OFF}+r_{ON}} + (f_{ON}(t=0) - \frac{r_{ON}}{r_{OFF}+r_{ON}})e^{-(r_{ON}+r_{OFF})t}$ and minimizing the χ^2 cost function, we determine that the inter-phenotype switching rates for the fast switchers are $r_{ON} = (0.0474 \pm 0.026) hours^{-1}$ and $r_{OFF} = (0.035 \pm 0.020) hours^{-1}$. The inter-phenotype switching rates for the slow switchers are: $r_{ON} = (0.0039 \pm 0.0002) hours^{-1}$ and $r_{OFF} = (0.007 \pm 0.0007) hours^{-1}$.

- blue line). When the cellular density exceeds a pre-set threshold (Figure 3-5b, upper panel - red line) a pump is activated (Figure 3-5b, middle panel) which dilutes the culture back below the threshold. By measuring the amount of media needed to maintain the culture in this way, we can accurately calculate the population's growth rate (Figure 3-5b, lower panel) over time.

To quantify the relationship between population diversity and growth rate, we measured the growth rate dynamics as the population transitions from one environment to the other. We first prepared separate cultures of fast and slow switchers by allowing them to reach a steady-state YFP expression distribution in environments E1 and E2 (Figs 3-3c-d, bottom panels). Subsequently, cells were transferred to the other environment (at time t = 0 as defined in Figs 3-5c-d) during which the population growth rate γ was monitored (Figs 3-5c-d). Figure 3-5c demonstrates the growth rate dynamics when the environment switches from E2 to E1. The growth rates of both the fast and slow switchers initially decreased rapidly within the first few hours since most cells were initially in the unfit (OFF) state and were unable to produce their own uracil. After a lag period of 3-5 hours, both fast and slow switching

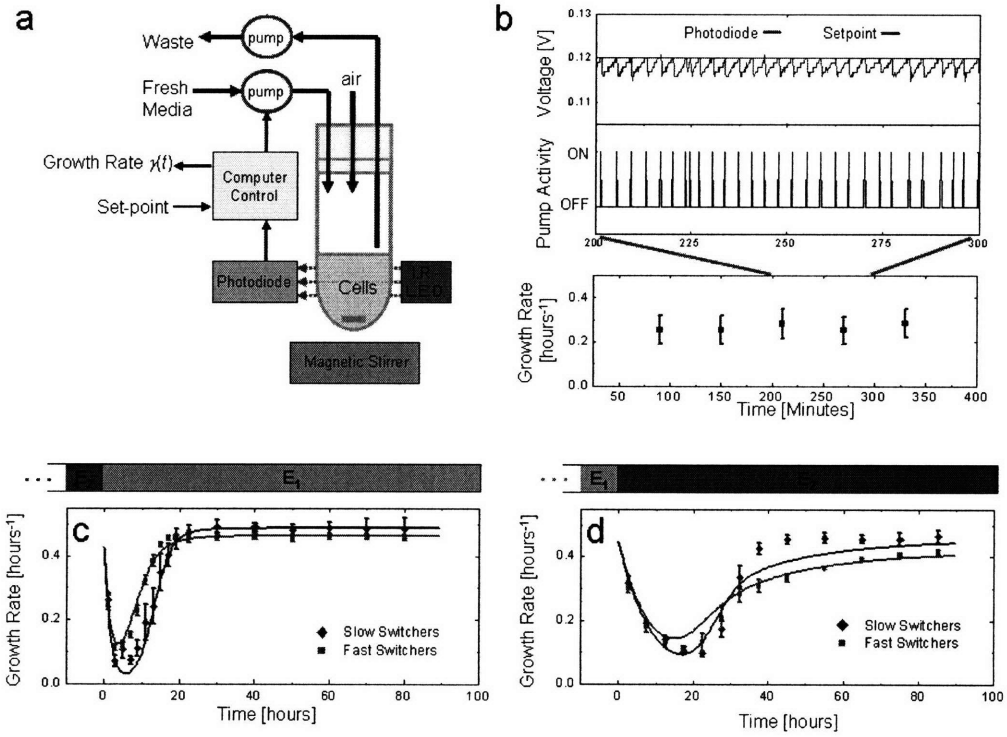

Figure 3-5: a, Schematic of the custom-made turbidostat setup used for all growth rate measurements. An infrared LED (dark gray) and photodiode (light gray) pair were used to continuously measure the relative optical density of the culture. A peristaltic pump intermittently provided fresh media to dilute the population while second continuously operating peristaltic pump coupled to a pickup tube kept the culture volume fixed. b, Whenever the photodiode voltage (blue line) went above a pre-set threshold (red line) the pump was activated (black line) to provide fresh media. The pumping rate was then used to calculate the dynamic population growth rate $\gamma(t)$ (blue circles). c, Growth rates corresponding to cells prepared in an E2 history and transferred to E1 at t = 0 show a transition period and a steady state region. Fast switching cells (red line) recover from the effect of environment change faster than slow switching cells (blue line) but have a lower steady-state growth rate. d, Growth rates for cells prepared in an E1 history and transferred to E2 at t = 0. The red line corresponds to fast switchers and the blue line to slow switchers. Solid lines are generated by the model.

populations began to increase their growth rates, and reached a steady-state growth rate after about 18 hours (Figure 3-5c, Figure 3-6).

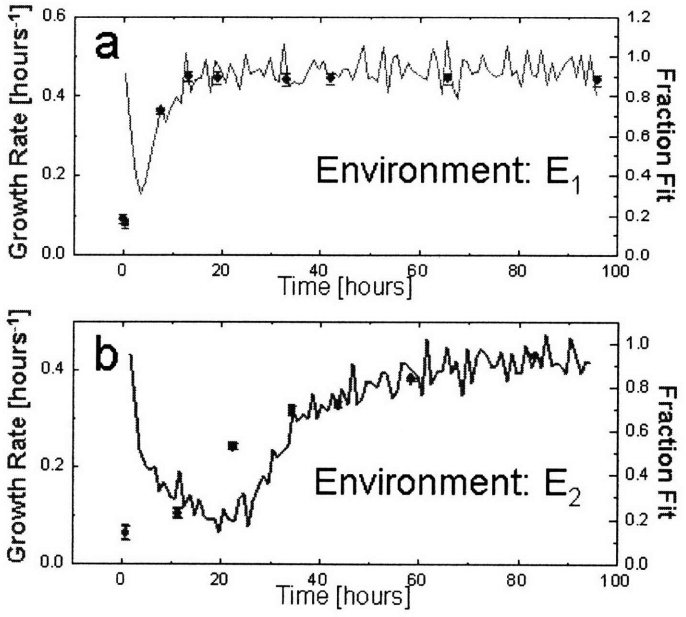

Figure 3-6: Dynamic measurement of cellular expression levels in the two environments. a, Fast switching cells growing in E2 were transferred to E1 at t = 0 and the fraction of ON and OFF cells were measured at several points in time (red symbols). The increase in the dynamic growth rate (gray line) coincides with the increase in the fraction of fit cells (ON). b, Fast switching cells growing in E1 were transferred to E2 at t = 0. The increase in growth rate coincides with a significant increase in the fraction of fit cells.

Although the growth rates for fast and slow switchers in Figures 3-5c-d show qualitatively similar dynamics, the quantitative differences between them illustrate the effects of diversity on the transient and steady-state growth rates. First, the fast switchers reached their steady-state about 4 hours sooner than the slow switchers. This is because they enter the new environment with a larger population of fit (ON) cells and also because more members of the initially large fraction of unfit (OFF) cells can switch to the fit (ON) state during the first few hours. During this initial phase the fast switching strategy is more competitive than the slow switching strategy. However, once the growth rates have reached steady-state, the slow switching strategy becomes

more competitive because the slow switchers less frequently transition to the unfit (OFF) phenotype. This results in a systematically larger steady-state growth rate compared to the fast switching strategy. We observe similar growth rate dynamics when cells transition from E1 to E2 (Figure 3-5d). The transient dynamics is overall slower in E2 most probably due to the slow synthesis of the toxic intermediate 5-Flurouracil. After the adaptation period in E2, both fast and slow switchers show similar growth rates to what was observed in environment E1 (Figure 3-5c). Taken together, this shows that after transitioning to a new environment, fast switchers have higher growth rates during the transition to steady-state but lower growth rates than the slow switchers during the steady-state.

3.7 A prediction for growth in periodic environments

To further demonstrate that the observed behavior is due to the differences in switching rates between fast and slow switchers, we built a quantitative model composed of ordinary differential equations. The model quantifies the number of cells in each of the two discrete phenotypes, ON and OFF, which both grow at different rates, γ_{ON}(t) and γ_{OFF}(t) respectively (Figure 3-1a). These growth rates reflect the fitness advantage or disadvantage conferred by the surrounding medium by quantifying both the steady state growth rate of each phenotype as well as the time required for each phenotype to reach steady state (Figure 3-7).

In addition, cells also transition from the OFF to ON (ON to OFF) phenotypes at a constant rate r_{ON} (r_{OFF}) as shown in Figure 3-4.

The model consists of two differential equations which characterize the dynamics of the number of cells in the ON and OFF states, N_{ON} and N_{OFF} respectively:

$$\frac{d}{dt}N_{ON} = \gamma_{ON}N_{ON} - r_{OFF}N_{ON} + r_{ON}N_{OFF}$$
$$\frac{d}{dt}N_{OFF} = \gamma_{OFF}N_{OFF} + r_{OFF}N_{ON} - r_{ON}N_{OFF}$$

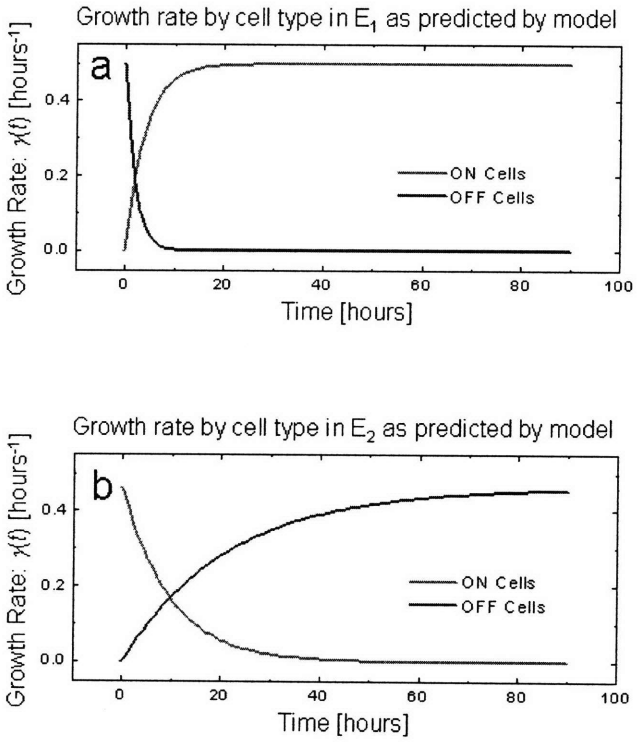

Figure 3-7: Modeling cellular growth in fluctuating environments. a, In environment E1, OFF cells (green) exponentially cease their growth rate due to the lack of intracellular uracil. Similarly, ON cells (orange) are assumed to need some time to recover from the previous environment (E2 with 5-FOA) and begin growing at their maximal rate. b, In E2, the opposite scenario occurs where ON cells (orange) decrease their growth rate with time as the effects of 5-FOA build up in the cell, and OFF cells (green) recover their maximal growth rate on a slightly longer timescale.

The population growth rate is given by:

$$\gamma(t) = \frac{\frac{d}{dt}N_{ON} + \frac{d}{dt}N_{OFF}}{N_{ON} + N_{OFF}}$$

The parameter r_{ON} (r_{OFF}) characterizes the rate of transitions to the ON (OFF) state. The values γ_{ON} and γ_{OFF} are the instantaneous growth rates of the ON and OFF phenotypes and depend on which environment the cells are in. In E1 the rates are given by $\gamma_{ON}(t) = \gamma_1(1 - e^{-r_1 t})$ and $\gamma_{OFF}(t) = \gamma_1 e^{-d_1 t}$, while in E2 the rates are given by $\gamma_{OFF}(t) = \gamma_2(1 - e^{-r_2 t})$ and $\gamma_{ON}(t) = \gamma_2 e^{-d_2 t}$ (Figure 3-7). In these equations t represents the amount of time since the last environmental transition, and the constants γ_1 and γ_2 represent the steady-state growth rates achieved by the fit phenotype in each environment. The constants r_1 and r_2 represent the amount of time it takes for cells to recover from a transition from their unfit to their fit environment. Similarly d_1 and d_2 represent the time it takes for unfit cells to cease growing after a transition from their fit environment to their unfit environment. The parameters r_1, r_2, d_1, and d_2 (Section 3.10.4) are fit to the data in Figures 3-5c-d by minimizing the χ^2 cost function with custom Matlab scripts. The model fit is depicted by the red and blue lines in Figures 3-5c-d. The only difference between fast and slow switching cells in the model is the difference between transition rates (r_{ON}, r_{OFF}), indicating that the observed growth rate differences between fast and slow switching cells can be accounted for by switching rates alone.

Using this population dynamics model, we can now predict which of the two switching strategies is more beneficial in fluctuating environments. In the model, the environment periodically alternates between environment E1 (with a duration T1) and E2 (with a duration T2) as depicted in Figure 3-1a. The fitness Γ of the population is defined as the average population growth rate over one period. The resulting fitness is calculated as a function of T1 and T2 for the fast (Γ_{fast}) and slow (Γ_{slow}) switching strategies. For both strategies short environmental periods lead to decreased fitness levels, since more time is spent in the transient recovery stage of growth. However, the fitness difference between slow switching and fast switching

populations can either be positive or negative depending on T1 and T2 (Figure 3-8a).

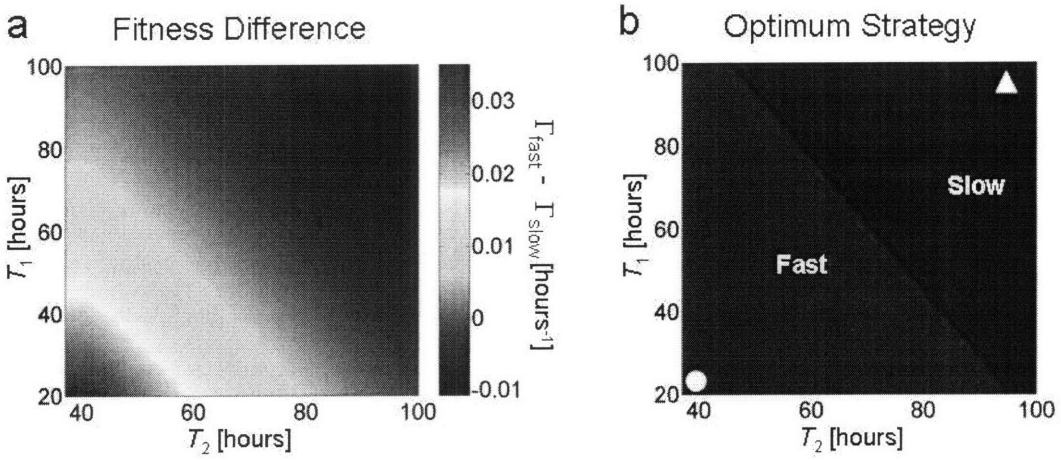

Figure 3-8: a, Heat map showing the predicted fitness difference (mean growth rate) between fast and slow switching cells as a function of the environmental period (T1, T2). For short periods, the fast switchers display higher mean growth rates. On the other hand, for longer periods slow switchers show a fitness advantage over the fast switchers. b, Phase diagram demonstrating regions where fast or slow switchers are predicted to be more fit. Two points were chosen in each part of the phase diagram to be further explored. The values of these points correspond to T1 = 20 hours, T2 = 37 hours (circle) and T1 = 96 hours, T2 = 96 hours (triangle).

This analysis shows that in rapidly changing environments fast switchers would out-compete the slow switchers, whereas for slowly varying environments the slow switchers would dominate. Figure 3-8b shows which of the two strategies is more beneficial given the times T1 and T2.

3.8 Growth rate dynamics in periodic environments

To experimentally test these predictions, we choose two points in the (T1, T2) space to measure fitness: T1 = 20 hours, T2 = 37 hours (circle, Figure 3-8b) and T1 = 96 hours, T2 = 96 hours (triangle, Figure 3-8b). As with the previous experiments, we grow the cells to steady state (4 days) in E2 then transition them to E1 at t = 0. After T1 hours, the extracellular environment was changed again to E2 for an additional T2 hours. The dynamically measured growth rates for these two fluctuating environments

are shown for both fast and slow switching populations in Figures 3-9a-b along with the population growth rates predicted by the model (Figures 3-9a-b, solid lines).

As the phase diagram suggests, the fast switching cells out-compete the slow switchers in the rapidly changing environment (Figure 3-9a), while the situation is reversed in the slowly changing environment (Figure 3-9b). This is shown as a fitness difference between the fast and slow switchers in Figures 3-9e-f as quantified by the average growth rate. Here, in the rapidly changing environment the difference between growth rates is 20 percent, while in the slowly changing environment it is 6 percent. Due to exponential growth these differences greatly magnify the relative competitiveness over even moderate time-scales. This out-competition is made dramatically obvious when comparing the number of cells that would have been produced in a non-limiting environment at the end of each run (Figures 3-9c-d).

3.9 Discussion

Our data suggest that bet-hedging constitutes a simple, yet effective, survival strategy to cope with fluctuating environments. Following this strategy, an isogenic population improves its fitness by promoting phenotypic diversity so that, at any given time, some of its members are prepared for an unforeseen environmental fluctuation. The diversity is introduced naturally through the stochastic nature of gene expression allowing isogenic populations to spread the risk by not putting "all of their eggs in one basket". Recent work suggested that cell-to-cell variability can have a large impact on the fitness of a population during times of adversity [76, 77]. Here we show that it is the frequency of the environmental fluctuations that constrains the inter-phenotype transition rates. In particular, a population optimizes its chances of survival in fluctuating environments by tuning the phenotypic switching rates with respect to the durations of environmental exposures. In this 'resonant' condition bet-hedging provides an effective survival strategy by 'blindy' anticipating environmental changes. This strategy could be utilized by cellular populations that are lacking dedicated signal transduction machinery for particular extracellular signals or when

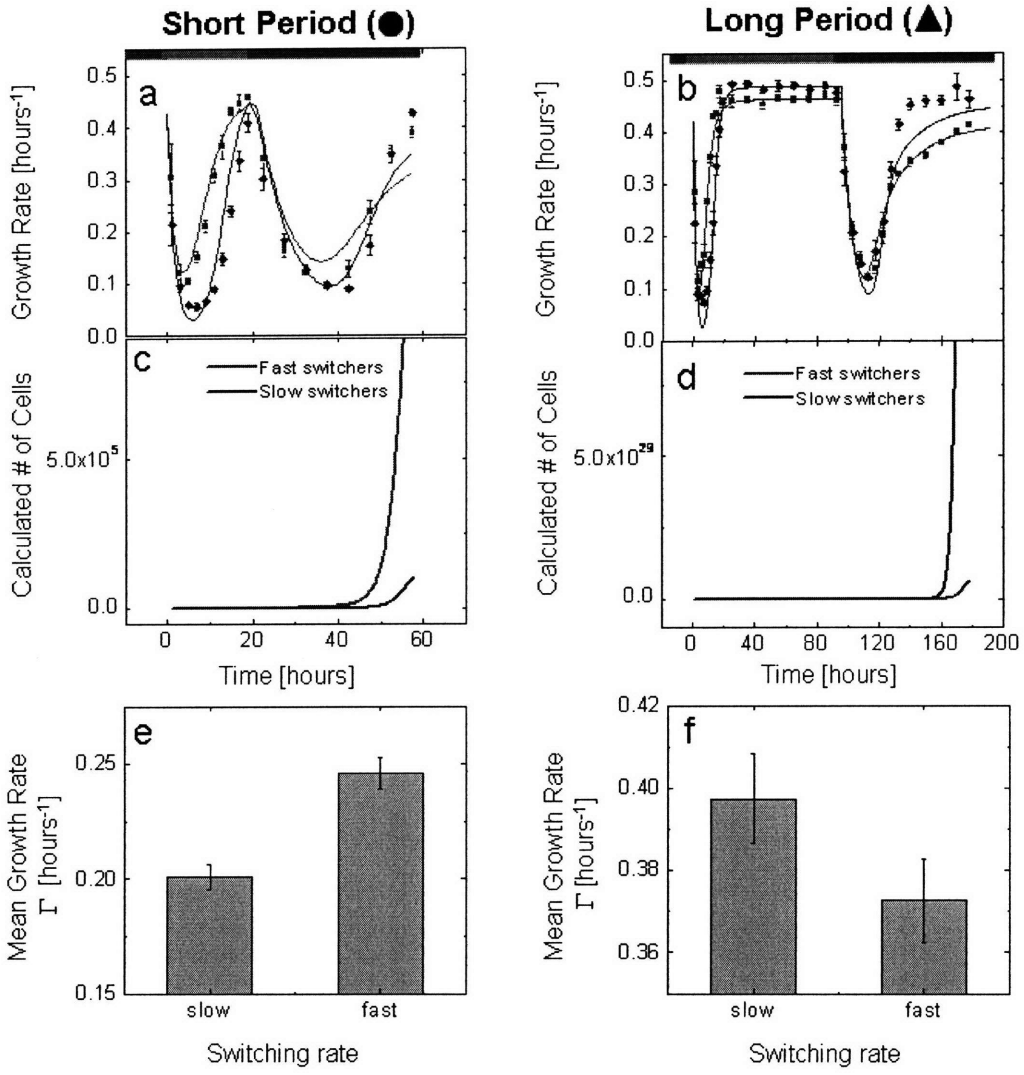

Figure 3-9: Testing the model predictions: growth dynamics in fluctuating environments with short and long periods. Growth rates for fast (red) and slow (blue) switchers grown for a, short environmental periods (20 hours in E1, 37 hours in E2, circle Figure 3-8b) and for b, long environmental periods (96 hours in E1, 96 hours in E2, triangle Figure 3-8b) are compared to the growth rates predicted by the model (solid lines). The calculated number of cells is shown for c, the short period and d, the long period environment, highlighting the relative fitness advantage of fast and slow switchers respectively. The experimentally measured fitness (mean growth rate) e, in the short period and f, the long period environment. Error bars represent a 2.8 percent error estimated from the standard deviation of the growth rate measurements.

it is crucial for a population to act at a faster timescale than is possible by signal transduction.

3.10 Methods

3.10.1 Measurement of inter-phenotypic switching rates

Cells were grown using non-selective synthetic dropout media (including 0.02 mg/ml uracil but lacking 5-FOA) in 2 percent raffinose. After overnight growth, cells were grown in 5 ml cultures for 30 hours in 2 percent raffinose ('raffinose history') and 2 percent raffinose and 2 percent galactose ('galactose history'). Subsequently, the raffinose and galactose history cells were separately grown for another 24 hours in media having 0.004 percent galactose and 0.00282μg/ml doxycycline (fast switchers) and 0.03 percent galactose and 0.0135μg/ml doxycycline (slow switchers). Culture volume during this period was 10 ml and the OD_{600} did not exceed 0.315. The expression distributions displayed were determined by flow cytometer (FACScan; Becton Dickinson).

We use a first order kinetic model to estimate the amount of time required for a cell to transition from one state to the other once it has decided to make the switch. Here we assume that ON cells produce YFP and Ura3 protein at a constant rate while OFF cells produce none. Removal of these proteins is assumed to be set by the dilution rate due to growth. Using this model we calculate that it requires approximately 1.3 hours for an average OFF cell to enter into the ON phenotype peak, while it requires approximately 6.5 hours for an ON cell to dilute enough YFP to be indistinguishable from an OFF cell. Both of these time-scales are smaller than the switching rates demonstrating that our approximation of the population as two distinct phenotypes is reasonable.

3.10.2 Strain construction and growth conditions

In order to couple the expression of the URA3 gene to the activity of the galactose pathway, the diploid strain MA0188 [5] was modified. The promoter of the URA3 gene was replaced by the GAL1 promoter using homologous recombination of transformed PCR product. The Candida albicans HIS5 gene was used as a marker. Integrations were verified by PCR.

Cultures were grown in synthetic dropout media with the appropriate amino-acid supplement and 2 percent raffinose as a carbon source. Prior to their turbidostat environments, cells were grown for 4 days in liquid culture in a shaker at 30C. The culture volume was 5 ml for the initial overnight growth and 10-15 ml afterwards. These 'pre-turbidostat' media contained uracil (at a final concentration of 0.02 mg/ml) and 5-FOA (at a final concentration of 0.19 mg/ml). During this period, cells were prevented from reaching stationary growth phase by serial dilution. Next, cells were washed with their prospective 'in-turbidostat' media (lacking uracil and 5-FOA) and were transferred to the turbidostat. The turbidostat maintains constant optical density levels (OD) corresponding to exponential growth ($0.05 < OD_{600} < 0.2$). When switching to environment E2, uracil (0.02 mg/ml) and 5-FOA (0.19 mg/ml) were added to the turbidostat media. The turbidostat temperature was maintained at 30°C. Fast and slow switching phenotypes were obtained by supplementing the synthetic dropout media with galactose and doxycycline. For the fast switching phenotype, 0.004 percent galactose and 0.00282μg/ml doxycycline were used. For the slow switching phenotype, 0.03 percent galactose and 0.0135μg/ml doxycycline were used. The expression distributions were determined by flow cytometer (FACScan; Becton Dickinson).

3.10.3 Turbidostat measurements

All growth rate measurements were made using a custom made turbidostat (Figs 3-5a-b). Cells were maintained at 10-15 ml volumes in test tubes while magnetic stir bars kept the cultures well mixed. An infrared LED/Photodiode pair (940 nm) was used

to measure the relative OD through a D/A converter and custom Labview software. A peristaltic pump provided fresh media to dilute the population whenever the OD went above a critical threshold value. A second continuously operating peristaltic pump coupled to a pickup tube kept the culture volume fixed.

Growth rates are calculated by recording the fraction of time the pump is actively providing fresh media during each hour-long interval and converting this into a pump flow rate by multiplying by the pump's maximal possible flow rate. This raw pump activity is then converted into a growth rate by normalizing with the volume of the turbidostat culture using the formula, $\gamma(t) = p(t)/V_{culture}$, where $p(t)$ represents the measured pumping rate in ml/hr and $V_{culture}$ represents the culture volume in ml.

3.10.4 Model parameters

The following table shows the parameters used in our model.

Parameter	Environment	Value
r_{ON} (slow)	1 & 2	0.0039 ± 0.0002 hours^{-1}
r_{OFF} (slow)	1 & 2	0.007 ± 0.0007 hours^{-1}
r_{ON} (fast)	1 & 2	0.0474 ± 0.026 hours^{-1}
r_{OFF} (fast)	1 & 2	0.035 ± 0.02 hours^{-1}
γ_1	1	0.50 ± 0.01 hours^{-1}
γ_2	2	0.46 ± 0.01 hours^{-1}
d_1	1	0.62 hours^{-1}
d_2	2	0.17 hours^{-1}
r_1	1	0.12 hours^{-1}
r_2	2	0.042 hours^{-1}

Chapter 4

Conclusion

Cells respond to environmental changes by activating their regulatory pathways. Regulatory pathways, especially in eukaryotic organisms, can have a complex net of interactions between the pathway components [23]. When it comes to understanding this complexity in cellular pathways, one approach to take would be to break apart their complex structure into modular pieces, understand the structure and function of the modules in detail, and at the end put the knowledge coming from each module back together to have a complete, 'big-picture' understanding of the original regulatory pathways. However, one crucial assumption is that biological networks are modular structures. Indeed, when we analyze the structure of cellular networks in detail [5,23], we often see that they are made up of recurrent, modular network structures, or network motifs. In this context, using the galactose signaling pathway of the budding yeast *Saccharomyces cerevisiae*, in the first half of my thesis, I have focused on understanding the effects of feedback loops in the pathway on the establishment and stability of cellular memory (the ability for cells to remember environmental conditions reliably for many generations).

The positive feedback loop established by Gal3p, the inducer of the galactose pathway, was responsible from building bistability into the wild type network. This positive feedback loop was providing cells with the ability to remember environmental conditions. Using energy landscapes to understand the dynamics of the system intuitively, we found that genetic noise was limiting the stability of cellular memory.

The feedback loop established by Gal80p (the repressor of the pathway) was shown to be decreasing the range of bistability in the wild-type network.

Cells in the wild are exposed to a variety of environmental conditions. Sometimes environments fluctuate frequently, but sometimes they tend to stay constant. Cells, on the other hand, often display multiple phenotypes and cellular populations can have different levels of phenotypic diversity as determined by the rates of phenotypic switching between the phenotypes displayed. Since each environment can impose its own fitness condition on a particular phenotype, an essential question arising in this context would be how the overall fitness of a population is optimized in terms of the frequency of environmental fluctuations and the level of phenotypic diversity.

In order to find an answer to this question, in the second half of my thesis, I focused on understanding the relationship between the rates of phenotypic switching between the two states of the bistable galactose network and the rates of fluctuations in the environmental conditions. The wild type galactose signaling pathway was re-engineered to give rise to varying population heterogeneity levels in terms of the relative number of cells expressed in their ON or OFF phenotypes. Cells were exposed to two different environments (each fit for one of the two phenotypes and unfit for the other phenotype) and population growth rates were measured dynamically when the cellular environment fluctuated. Our results showed that, in order to reach an optimum population fitness level, the frequency of environmental fluctuations should match the frequency of inter-phenotype switching.

The work composing my thesis demonstrates that systematically perturbing the underlying feedback loops in a network offers a general, powerful analysis method that can be applied to more complicated cellular networks in higher organisms to dissect the structure-function relationships of their components.

When we look at the types of network motifs recurrent in cellular pathways [3,18], we find that they are surprisingly small in number. In principle, there could be many more types of interaction topologies between the elements of endogenous pathways, but cells utilize only a small fraction of all possible interaction schemes in achieving desired pathway activity levels. Why are there only a small number of network motif

types in cellular pathways? How can environmental conditions affect the structure and recurrence frequency of network motifs? Future work focusing on dissecting the mechanisms which have been shaping the structure and function of gene networks will take us one step closer to understanding life at the cellular level.

Appendix A

Strain List

1. MA0182 = *MAT***a**/α, *ura3/ura3::URA3*-P_{TETO2}-*GAL3*, *his3::HIS3/his3*, *ade2::ADE2*-P_{MYO2}-*rtTA/ade2::ADE2*-P_{GAL1}-*YFP*, *gal3*Δ*::KanMX/gal3*Δ*::KanMX*

2. MA0188 = *MAT***a**/α, *ura3/ura3::URA3*-P_{TETO2}-*GAL80*, *his3::HIS3/his3*, *ade2::ADE2*-P_{MYO2}-*rtTA/ade2::ADE2*-P_{GAL1}-*YFP*, *gal80*Δ*::KanMX/gal80*Δ*::KanMX*

3. MA0207 = *MAT***a**/α, *his3::HIS3/his3*, *ade2/ade2::ADE2*-P_{GAL1}-*YFP*

4. MA0208 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2/ade2::ADE2*-P_{GAL2}-*YFP*

5. MA0210 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2/ade2::ADE2*-P_{GAL4}-*YFP*

6. MA0211 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2/ade2::ADE2*-P_{GAL6}-*YFP*

7. MA0212 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2/ade2::ADE2*-P_{GAL7}-*YFP*

8. MA0213 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2/ade2::ADE2*-P_{GAL10}-*YFP*

9. MA0215 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2::ADE2*-P_{MYO2}-*rtTA/ade2::ADE2*-P_{GAL1}-*YFP*, *gal2*Δ*::KanMX/gal2*Δ*::KanMX*

10. MA0226 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2::ADE2*-P_{MYO2}-*rtTA/ade2::ADE2*-P_{GAL1}-*YFP*, *gal3*Δ*::KanMX/gal3*Δ*::KanMX*

11. MA0231 = *MAT***a**/α, *ura3/ura3::URA3*, *his3::HIS3/his3*, *ade2::ADE2*-P_{GAL3}-

CFP/ade2::ADE2-P_{GAL1}*-YFP*

12. MA0239 = *MAT***a**/α, *ura3/ura3::URA3-*P_{TETO2}*-GAL80, his3::HIS3/his3, ade2::ADE2-*P_{MYO2}*-rtTA/ade2::ADE2-*P_{GAL1}*-YFP,trp1::TRP1-*P_{GAL3}*-CFP/trp1, gal80*Δ*::KanMX/gal80*Δ*::KanMX*

13. MA0242 = *MAT***a**/α, *ura3/ura3::URA3, his3::HIS3/his3, ade2::ADE2-*P_{GAL80}*-CFP/ade2::ADE2-*P_{GAL1}*-YFP*

14. MA0273 = *MAT***a**/α, *ura3/ura3::URA3-*P_{TETO2}*-GAL3-CFP, his3::HIS3/his3, ade2::ADE2-*P_{MYO2}*-rtTA/ade2::ADE2-*P_{GAL1}*-YFP*

15. MA0282 = *MAT***a**/α, *ura3/ura3::URA3-*P_{GAL3}*-GAL3-CFP, his3::HIS3/his3, ade2::ADE2-*P_{MYO2}*-rtTA/ade2::ADE2-*P_{GAL1}*-YFP, gal3*Δ*::KanMX/GAL3*

16. MA0283 = *MAT***a**/α, *ura3/ura3::URA3-*P_{GAL80}*-GAL80-CFP, his3::HIS3/his3, ade2::ADE2-*P_{MYO2}*-rtTA/ade2::ADE2-*P_{GAL1}*-YFP, GAL80/gal80*Δ*::KanMX*

17. MA0291 = *MAT***a**/α, *ura3/ura3::URA3-*P_{TETO2}*-GAL80-CFP, his3::HIS3/his3, ade2::ADE2-*P_{MYO2}*-rtTA/ade2::ADE2-*P_{GAL1}*-YFP*

Appendix B

Galactose Depletion Experiments

Construct	Galactose consumption rate c (mM·(OD_{600}·hours)$^{-1}$)	Density of culture when galactose concentration changes by 10% (OD_{600})†
WT	0.084±0.010	0.431
gal80Δ P_{TET}*GAL80*	0.062±0.041	0.574
gal3Δ P_{TET}*GAL3*	0.639±0.083	0.056

1

†The density of the culture was calculated from kg/c, where k is cell division rate, g is the change in galactose concentration. g=0.002 percent, which corresponds to a 10 percent change when the initial galactose concentration is 0.020 percent. At the beginning of the experiment $OD_{600} = 10^{-7}$. A concentration of 30μg/ml doxycycline was used to obtain a maximum *GAL3* and *GAL80* expression.

Bibliography

[1] W.S. Klug and M.R. Cummings. *Concepts of Genetics, 7th Edition.* Prentice Hall, Upper Saddle River, New Jersey, 2003.

[2] M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[3] T.I. Lee et al. Transcriptional regulatory networks in saccharomyces cerevisiae. *Science*, 298(5594):799–804, 2002.

[4] S. Mangan and U. Alon. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences*, 100(21):11980–11985, 2003.

[5] M. Acar, A. Becskei, and A. van Oudenaarden. Enhancement of cellular memory by reducing stochastic transitions. *Nature*, 435(7039):228–32, 2005.

[6] M. Acar, J.T. Mettetal, and A. van Oudenaarden. Phenotypic bet-hedging as a survial strategy in yeast populations. *in submission.*

[7] E.M. Ozbudak, M. Thattai, I. Kurtser, A.D. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature Genetics*, 31(1):69–73, 2002.

[8] J. Paulsson. Summing up the noise in gene networks. *Nature*, 427:415–418, 2004.

[9] M. Kaern, T.C. Elston, W.J. Blake, and J.J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet*, 6(6):451–64, 2005.

[10] J.M. Raser and E.K. O'Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309(5743):2010–2013, 2005.

[11] W.J. Blake, M. KAern, C.R. Cantor, and J.J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, 2003.

[12] J.M. Raser and E.K. O'Shea. Control of stochasticity in eukaryotic gene expression. *Science*, 304(5678):1811–1814, 2004.

[13] M.F. Wernet, E.O. Mazzoni, A. Celik, D.M. Duncan, I. Duncan, and C. Desplan. Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature*, 440(7081):174–180, 2006.

[14] M. Thattai and A. van Oudenaarden. Stochastic gene expression in fluctuating environments. *Genetics*, 167(1):523–530, 2004.

[15] E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–2078, 2005.

[16] DM Wolf, VV Vazirani, and AP Arkin. Diversity in times of adversity: probabilistic strategies in microbial survival games. *J Theor Biol*, 234(2):227–53, 2005.

[17] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[18] S.S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. Network motifs in the transcriptional regulation network of escherichia coli. *Nature Genetics*, 31(1):64–68, 2002.

[19] A. Becskei, B. Seraphin, and L. Serrano. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *The EMBO Journal*, 20:2528–2535, 2001.

[20] J.E. Ferrell. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr. Opin. Chem. Biol*, 6:140–148, 2002.

[21] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–593, 2000.

[22] M. Freeman. Feedback control of intercellular signalling in development. *Nature*, 408:313–319, 2000.

[23] E.H. Davidson et al. A genomic regulatory network for development. *Science*, 295:1669–1678, 2002.

[24] H. Xie, M. Ye, R. Feng, and T. Graf. Stepwise reprogramming of b cells into macrophages. *Cell*, 117(5):663–676, 2004.

[25] W. Xiong. A positive-feedback-based bistable'memory module'that governs a cell fate decision. *Nature*, 426:460–465, 2003.

[26] B.M. Shykind, S.C. Rohani, S. O'Donnell, A. Nemes, M. Mendelsohn, Y. Sun, R. Axel, and G. Barnea. Gene switching and the stability of odorant receptor gene choice. *Cell*, 117(6):801–815, 2004.

[27] N.I. Markevich, J.B. Hoek, and B.N. Kholodenko. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J. Cell Biol.*, 164(3):353–359, 2004.

[28] T.S. Gardner, C.R. Cantor, and J.J. Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403(6767):339–342, 2000.

[29] M.R. Atkinson, M.A. Savageau, J.T. Myers, and A.J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in escherichia coli. *Cell*, 113(5):597–607, 2003.

[30] F.J. Isaacs, J. Hasty, C.R. Cantor, and J.J. Collins. Prediction and measurement of an autoregulatory genetic module. *Proceedings of the National Academy of Sciences*, 100(13):7714–7719, 2003.

[31] D. Angeli, J.E. Ferrell, and E.D. Sontag. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences*, 101(7):1822–1827, 2004.

[32] K. Lai, M.J. Robertson, and D.V. Schaffer. The sonic hedgehog signaling system as a bistable genetic switch. *Biophysical Journal*, 86:2748–2757, 2004.

[33] S.R. Biggar and G.R. Crabtree. Cell signaling can direct either binary or graded transcriptional responses. *The EMBO Journal*, 20(12):3167–3176, 2001.

[34] T. Ideker, V. Thorsson, J.A. Ranish, R. Christmas, J. Buhler, J.K. Eng, R. Bumgarner, D.R. Goodlett, R. Aebersold, and L. Hood. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, 292(5518):929–934, 2001.

[35] T. Suzuki-Fujimoto, M. Fukuma, KI Yano, H. Sakurai, A. Vonika, SA Johnston, and T. Fukasawa. Analysis of the galactose signal transduction pathway in saccharomyces cerevisiae: interaction between gal3p and gal80p. *Molecular and Cellular Biology*, 16(5):2504–2508, 1996.

[36] D.J. Timson, H.C. Ross, and R.J. Reece. Gal3p and gal1p interact with the transcriptional repressor gal80p to form a complex of 1: 1 stoichiometry. *Biochem. J*, 363:515–520, 2002.

[37] G. Peng and J.E. Hopper. Gene activation by interaction of an inhibitor with a cytoplasmic signaling protein. *Proceedings of the National Academy of Sciences*, 99(13):8548, 2002.

[38] A. Mizutani and M. Tanaka. Regions of gal4 critical for binding to a promoter in vivo revealed by a visual dna-binding analysis. *EMBO Journal*, 22:2178–2187, 2003.

[39] M. Johnston, J.S. Flick, and T. Pexton. Multiple mechanisms provide rapid and stringent glucose repression of gal gene expression in saccharomyces cerevisiae. *Molecular and Cellular Biology*, 14(6):3834–3841, 1994.

[40] R. Wieczorke, S. Krampe, T. Weierstall, K. Freidel, CP Hollenberg, and E. Boles. Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in saccharomyces cerevisiae. *FEBS Letters*, 464(3):123–128, 1999.

[41] E.M. Ozbudak, M. Thattai, H.N. Lim, B.I. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilization network of escherichia coli. *Nature*, 427:737–740, 2004.

[42] PJ Bhat and JE Hopper. The mechanism of inducer formation in gal3 mutants of the yeast galactose system is independent of normal galactose metabolism and mitochondrial respiratory function. *Genetics*, 128(2):233–239, 1991.

[43] J.R. Rohde, J. Trinh, and I. Sadowski. Multiple signals regulate gal transcription in yeast. *Molecular and Cellular Biology*, 20(11):3880–3886, 2000.

[44] J. Hasty, J. Pradines, M. Dolnik, and J.J. Collins. Noise-based switches and amplifiers for gene expression. *Proceedings of the National Academy of Sciences*, 97(5):2075–2080, 2000.

[45] W. Bialek. Stability and noise in biochemical switches. *Arxiv preprint cond-mat/0005235*, 2000.

[46] T.B. Kepler and T.C. Elston. Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophysical Journal*, 81(6):3116–3136, 2001.

[47] HA Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[48] G. Peng and J.E. Hopper. Evidence for gal3p's cytoplasmic location and gal80p's dual cytoplasmic-nuclear location implicates new mechanisms for controlling

gal4p activity in saccharomyces cerevisiae. *Molecular and Cellular Biology*, 20(14):5140–5148, 2000.

[49] K. Melcher and H.E. Xu. Gal80- gal80 interaction on adjacent gal4p binding sites is required for complete gal gene repression. *The EMBO Journal*, 20:841–851, 2001.

[50] M. Verma, P.J. Bhat, and KV Venkatesh. Quantitative analysis of gal genetic switch of saccharomyces cerevisiae reveals that nucleocytoplasmic shuttling of gal80p results in a highly sensitive response to galactose. *Journal of Biological Chemistry*, 278(49):48764–48769, 2003.

[51] B. Ren et al. Genome-wide location and function of dna binding proteins. *Science*, 290(5500):2306–2309, 2000.

[52] S. Urlinger, U. Baron, M. Thellmann, M.T. Hasan, H. Bujard, and W. Hillen. Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proceedings of the National Academy of Sciences*, 97(14):7963–7968, 2000.

[53] HH Sillje, EG ter Schure, AJ Rommens, PG Huls, CL Woldringh, AJ Verkleij, J. Boonstra, and CT Verrips. Effects of different carbon fluxes on g1 phase duration, cyclin expression, and reserve carbohydrate metabolism in saccharomyces cerevisiae. *Journal of Bacteriology*, 179(21):6560–6565, 1997.

[54] D. Cohen. Optimizing reproduction in a randomly varying environment. *J Theor Biol*, 12(1):119–29, 1966.

[55] R. Levins. *Evolution in Changing Environments: Some Theoretical Explorations(Mpb-2)*. Princeton University Press, 1968.

[56] W.M. Schaffer. Optimal reproductive effort in fluctuating environments. *The American Naturalist*, 108(964):783–790, 1974.

[57] S.C. Stearns. Life-history tactics: A review of the ideas. *The Quarterly Review of Biology*, 51(1):3–47, 1976.

[58] J. Seger and H.J. Brockmann. What is bet-hedging. *Oxford Surveys in Evolutionary Biology*, 4:182–211, 1987.

[59] M. Slatkin. Hedging one's evolutionary bets. *Nature*, 250(5469):704–705, 1974.

[60] M.W. van der Woude and A.J. Baumler. Phase and antigenic variation in bacteria. *Clinical Microbiology Reviews*, 17(3):581–611, 2003.

[61] A. Hernday, M. Krabbe, B. Braaten, and D. Low. Self-perpetuating epigenetic pili switches in bacteria. *Proceedings of the National Academy of Sciences*, 99:16470–16476, 2002.

[62] D.M. Wolf and A.P. Arkin. Fifteen minutes of fim: Control of type 1 pili expression in e. coli. *OMICS: A Journal of Integrative Biology*, 6(1):91–114, 2002.

[63] P.N. Danese, L.A. Pratt, S.L. Dove, and R. Kolter. The outer membrane protein, antigen 43, mediates cell-to-cell interactions within escherichia coli biofilms. *Molecular Microbiology*, 37(2):424–432, 2000.

[64] AD Grossman. Genetic networks controlling the initiation of sporulation and the development of genetic competence in bacillus subtilis. *Annu Rev Genet*, 29:477–508, 1995.

[65] G.M. Suel, J. Garcia-Ojalvo, L.M. Liberman, and M.B. Elowitz. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature*, 440(7083):545–550, 2006.

[66] A. Novick and M. Weiner. Enzyme induction as an all-or-none phenomenon. *Proceedings of the National Academy of Sciences*, 43(7):553–566, 1957.

[67] N. Dhar and J.D. McKinney. Microbial phenotypic heterogeneity and antibiotic tolerance. *Curr Opin Microbiol*, 10(1):30–38, 2007.

[68] E. Kussell, R. Kishony, N.Q. Balaban, and S. Leibler. Bacterial persistence a model of survival in changing environments. *Genetics*, 169(4):1807–1814, 2005.

[69] M. Lachmann and E. Jablonka. The inheritance of phenotypes: An adaptation to fluctuating environments. *Journal of Theoretical Biology*, 181(1):1–9, 1996.

[70] B.B. Kaufmann and A. van Oudenaarden. Stochastic gene expression: from single molecules to the proteome. *Curr Opin Genet Dev*, 7(2):107–12, 2007.

[71] D. Dubnau and R. Losick. Bistability in bacteria. *Mol. Microbiol*, 61:564–72, 2006.

[72] J. Hasty, D. McMillen, and J.J. Collins. Engineered gene circuits. *Nature*, 420(6912):224–230, 2002.

[73] K.M. Hawkins and C.D. Smolke. The regulatory roles of the galactose permease and kinase in the induction response of the gal network in saccharomyces cerevisiae. *J Biol Chem*, 281(19):13485–13492, 2006.

[74] J.D. Boeke, F. Croute, and G.R. Fink. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Molecular Genetics and Genomics*, 197(2):345–346, 1984.

[75] V. Bryson and W. Szybalski. Microbial selection. *Science*, 115(3003):45–51, 1952.

[76] A.L. Bishop, F.A. Rab, E.R. Sumner, and S.V. Avery. Phenotypic heterogeneity can enhance rare-cell survival in 'stress sensitive'yeast populations. *Molecular Microbiology*, 63(2):507–520, 2007.

[77] W.J. Blake, G. Balazsi, M.A. Kohanski, F.J. Isaacs, K.F. Murphy, Y. Kuang, C.R. Cantor, D.R. Walt, and J.J. Collins. Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular Cell*, 24(6):853–865, 2006.